

neurons responded sluggishly to stimulation of a dopamine receptor known as D1. The mice had an altered distribution pattern of the G protein-coupled receptor kinase 2 protein, GRK2, which decreases D1's activity. FMRP and GRK2 were also found to bind to one another, and blocking GRK2 with a drug corrected the mice's D1 signalling deficiency. Drugs that boosted D1 signalling also restored some behaviour deficits in the mice. Zhuo and his colleagues conclude that FMRP is a player in the dopamine signalling architecture of the forebrain and that this may explain some of the problems caused by its deficit in fragile X syndrome.

CHEMISTRY

Ro-taxing synthesis

Angew. Chem. Int. Edn doi: 10.1002/anie.200803056 (2008)

Chemists are fascinated by rotaxanes — a class of compound made from a dumbbell-shaped chemical inside a molecular ring — because such structures could be useful in making microscale machines. Sheng-Hsien Chiu and his colleagues at the National Taiwan University have now created the smallest rotaxane to date (pictured right), made from just 76 atoms, with a molecular mass of 510 daltons.

Their ring is a crown ether — a circle of alternating oxygen atoms and ethylene groups — and their dumbbell's axle is a dialkylammonium ion. After threading the axle through the ring, grinding the mixture up with a tetrazine caps the axle with pyridazine groups to make the rotaxane.

The scientists hope that they can shave a few atoms off the rotaxane to beat their own record.

EVOLUTIONARY BIOLOGY

Eggs in many baskets

Proc. R. Soc. Lond. B doi: 10.1098/rspb.2008.0794 (2008)

In Bibron's toadlet (*Pseudophryne bibronii*) males are the nestmakers and females shop around. Females of this Australian amphibian lay their eggs in the nests of up to eight males. This makes it the first frog or toad in which such 'sequential polyandry'

has been genetically proven, and the most extreme example of this mating tactic seen in any vertebrate.

Using DNA analysis and observations, Phillip Byrne, of Monash University in Melbourne, and his colleague showed that females that laid eggs in more nests had more surviving tadpoles. They suggest that polyandry is an insurance policy against nests flooding or drying out, and that other egg-layers in which dad tends the brood in a risky environment may do the same.

FLUID DYNAMICS

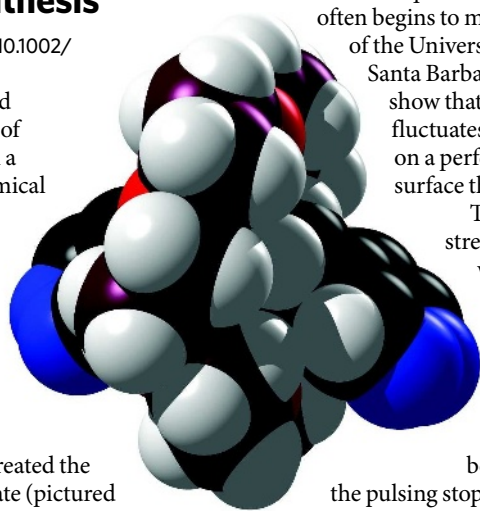
Making water wander

Phys. Rev. Lett. **101**, 114501 (2008)

As a stream of liquid flows down a slope, it often begins to meander. Björn Birnir of the University of California, Santa Barbara, and his colleagues show that if the liquid flow fluctuates this can happen even on a perfectly smooth, clean surface that does not erode.

The researchers ran a stream of a water mixed with glycerine down a plastic sheet. As the flow changed from smooth to pulsed, the rivulet began meandering. The wavy path can be sustained once

the pulsing stops, as the stream hits droplets left on the surface during its earlier wanderings.



ATMOSPHERIC CHEMISTRY

Glasses in the sky

Atmos. Chem. Phys. **8**, 5221–5244 (2008)

Some solutions of water-soluble organic compounds form glasses below -43°C . The upper troposphere, where cirrus clouds form, is this cold; glass formation here could impede the water uptake of aerosol particles, ice nucleation and crystal growth, with implications for cloud formation.

Thomas Koop of Bielefeld University in Germany and his colleagues investigated glass formation in various solutions under realistic atmospheric conditions in the lab. They found that aerosol particles enriched with large organic molecules are most likely to form glasses at low temperatures and high relative humidity. Modelling studies should be used to assess the regional and global impact of atmospheric glass formation, say the authors.

JOURNAL CLUB

Nicola Hamilton and David Attwell
University College London

Two neuroscientists are surprised by the link between a brain-chemical transporter and sexual orientation.

Many nerve cells in the brain release the chemical neurotransmitter glutamate to signal to other neurons via receptors. Dedicated transporters then remove glutamate from the extracellular space to end signalling.

Cystine-glutamate exchangers are unusual glutamate transporters because they do the reverse, adding glutamate to the extracellular space while removing cystine. David Featherstone of the University of Illinois, Chicago, and his colleagues have found that in the fruitfly *Drosophila melanogaster*, knocking down expression of a cystine-glutamate exchanger in non-neuronal glial cells leads to a dramatic change in the sexual behaviour of male flies: they mate with both males and females owing to altered processing of sex-specific chemosensory cues (Y. Grosjean *et al. Nature Neurosci.* **11**, 54–61; 2008).

This behaviour may be caused by an increase in the number of glutamate signalling receptors, which is induced by the fall in extracellular glutamate concentration that follows transporter knockdown. Indeed, the effect of the knockdown could be reversed by feeding the flies a drug that reduces glutamate signalling, and could be mimicked by feeding normal flies a drug that enhances glutamate signalling.

These studies raise questions about whether human sexual orientation, long assumed to be due to a mix of genes and environment, could also be altered by perturbations of neurotransmitter signalling. Could differences in such signalling contribute to different sexual preferences?

The possibility of altering sexual preference pharmacologically is worrying. We cannot rule out a future regression to the twentieth-century idea that sexual behaviour should be regulated by society.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

The other North American election

As Canadian scientists work to maintain their international reputation, a little encouragement from the election candidates would be appreciated.

Canadians go to the polls to elect a new government on 14 October. Although the initial stages of the campaign focused on the environment (see page 268), the two major parties, Conservative and Liberal, have said little or nothing about science policy in general. This is a shame. Canada saw big boosts to its research funding from the late 1990s to 2000, including the creation of Canada Research Chairs, which brought good people into the country, and the Canada Foundation for Innovation, which pumped billions into infrastructure. Those investments have been maintained, and science funding is still on the rise. But the gains are vulnerable in a competitive international market, warns the prime minister's former science adviser, Arthur Carty: "We have to be careful, having reached the top of the mountain, that we don't slide down the other side very quickly."

Both parties promise to provide financial incentives to innovative companies, especially those involving green technologies. But broader questions of research funding have so far not come up in the campaign. This may reflect a lack of difference between the parties on the issue, or perhaps just a lack of urgency; with the exception of climate change, the general mood on science policy seems to be 'if it ain't broke, don't fix it'.

But many argue that it is broke. The retired right-wing politician Preston Manning slammed the system in the Canadian media last December after the recent shortage of nuclear isotopes. He lambasted Canada for its lack of a federal science department or ministry and

the dearth of scientific or engineering training among parliamentarians. The office of the National Science Advisor was abolished earlier this year when Carty stepped down (see *Nature* 451, 505; 2008). And the committee that now advises the prime minister on matters of science is packed with industrial as well as scientific experts.

Indeed, many Canadian scientists are seeing, and complaining about, an undue emphasis on commercially focused research over long-term basic research. Such complaints are heard in many other countries too. But in Canada the problem is compounded by the fact that the current government has channelled new science funds into four restrictive priority areas

— natural resources, environment, health and information technology — and that scientists are often required to scrounge matching funds from elsewhere to top up their grants. Furthermore, the government this month defined sub-priority areas that mix in obvious commercial influences: alongside 'Arctic monitoring', for example, sits 'energy production from the oil sands'.

The Canadian election's focus on climate, at least, is welcome. But one always hopes that research funding will warrant a mention in political manifestos. That hasn't happened yet in Canada — and it should. ■

"Many Canadian scientists are complaining about an undue emphasis on commercially focused research over long-term basic research."

Handle with care

Ecologists must research how best to intervene in and preserve ecosystems.

For many people — including many scientists — 'nature' is defined by a negative: it exists where people do not. Nature lies outside the urban and agricultural realms, in regions of Earth where natural processes are unimpeded. Nature is where fallen logs rot and acorns grow, wildfires turn woodlands into meadows, and barrier islands shift with the currents — all without human interference. By extension, this definition suggests that nature is best protected by keeping humans far away, so that it can continue to run itself.

But there is a serious problem with this view. If nature is defined as a landscape uninfluenced by humankind, then there is no nature on the planet at all. Prehistoric peoples changed their surrounding ecosystems, whether by installing orchards in the Amazon or — according to one increasingly accepted theory — by hunting many large mammals to extinction in North America. And modern humans are changing the global environment even more profoundly, whether

through planet-wide climate change, or by the worldwide movement of synthetic chemicals through the food chain. Today there is no place untouched by man — a point made by environmentalist Bill McKibben as early as 1989 in the starkly titled *The End of Nature*.

Nature doesn't have to end if we stop defining it by humankind's absence. Humans prize natural spaces because they are historic, culturally significant, aesthetic and scientifically interesting — and, increasingly, because they have been recognized as providing essential services such as filtering water, ameliorating storm surge, providing fish, game and timber, and sequestering carbon. Ecosystems that are valuable for one or more of these reasons can be identified by quantifiable biological traits, such as the presence of certain key species or processes. In the Białowieża forest of eastern Europe, which has a long history of human activity, for example, one could cite the presence of European bison and of a large amount of dead wood as characteristics worth preserving (see page 277).

Retaining such characteristics takes more than the absence of active destruction. It is precisely because of humanity's pervasive influence that even the least changed ecosystems need help surviving in the future. Białowieża's core is so small that the dynamic processes that once drove its mosaic of different micro-ecosystems probably can't

The other North American election

As Canadian scientists work to maintain their international reputation, a little encouragement from the election candidates would be appreciated.

Canadians go to the polls to elect a new government on 14 October. Although the initial stages of the campaign focused on the environment (see page 268), the two major parties, Conservative and Liberal, have said little or nothing about science policy in general. This is a shame. Canada saw big boosts to its research funding from the late 1990s to 2000, including the creation of Canada Research Chairs, which brought good people into the country, and the Canada Foundation for Innovation, which pumped billions into infrastructure. Those investments have been maintained, and science funding is still on the rise. But the gains are vulnerable in a competitive international market, warns the prime minister's former science adviser, Arthur Carty: "We have to be careful, having reached the top of the mountain, that we don't slide down the other side very quickly."

Both parties promise to provide financial incentives to innovative companies, especially those involving green technologies. But broader questions of research funding have so far not come up in the campaign. This may reflect a lack of difference between the parties on the issue, or perhaps just a lack of urgency; with the exception of climate change, the general mood on science policy seems to be 'if it ain't broke, don't fix it'.

But many argue that it is broke. The retired right-wing politician Preston Manning slammed the system in the Canadian media last December after the recent shortage of nuclear isotopes. He lambasted Canada for its lack of a federal science department or ministry and

the dearth of scientific or engineering training among parliamentarians. The office of the National Science Advisor was abolished earlier this year when Carty stepped down (see *Nature* 451, 505; 2008). And the committee that now advises the prime minister on matters of science is packed with industrial as well as scientific experts.

Indeed, many Canadian scientists are seeing, and complaining about, an undue emphasis on commercially focused research over long-term basic research. Such complaints are heard in many other countries too. But in Canada the problem is compounded by the fact that the current government has channelled new science funds into four restrictive priority areas

— natural resources, environment, health and information technology — and that scientists are often required to scrounge matching funds from elsewhere to top up their grants. Furthermore, the government this month defined sub-priority areas that mix in obvious commercial influences: alongside 'Arctic monitoring', for example, sits 'energy production from the oil sands'.

The Canadian election's focus on climate, at least, is welcome. But one always hopes that research funding will warrant a mention in political manifestos. That hasn't happened yet in Canada — and it should. ■

"Many Canadian scientists are complaining about an undue emphasis on commercially focused research over long-term basic research."

Handle with care

Ecologists must research how best to intervene in and preserve ecosystems.

For many people — including many scientists — 'nature' is defined by a negative: it exists where people do not. Nature lies outside the urban and agricultural realms, in regions of Earth where natural processes are unimpeded. Nature is where fallen logs rot and acorns grow, wildfires turn woodlands into meadows, and barrier islands shift with the currents — all without human interference. By extension, this definition suggests that nature is best protected by keeping humans far away, so that it can continue to run itself.

But there is a serious problem with this view. If nature is defined as a landscape uninfluenced by humankind, then there is no nature on the planet at all. Prehistoric peoples changed their surrounding ecosystems, whether by installing orchards in the Amazon or — according to one increasingly accepted theory — by hunting many large mammals to extinction in North America. And modern humans are changing the global environment even more profoundly, whether

through planet-wide climate change, or by the worldwide movement of synthetic chemicals through the food chain. Today there is no place untouched by man — a point made by environmentalist Bill McKibben as early as 1989 in the starkly titled *The End of Nature*.

Nature doesn't have to end if we stop defining it by humankind's absence. Humans prize natural spaces because they are historic, culturally significant, aesthetic and scientifically interesting — and, increasingly, because they have been recognized as providing essential services such as filtering water, ameliorating storm surge, providing fish, game and timber, and sequestering carbon. Ecosystems that are valuable for one or more of these reasons can be identified by quantifiable biological traits, such as the presence of certain key species or processes. In the Białowieża forest of eastern Europe, which has a long history of human activity, for example, one could cite the presence of European bison and of a large amount of dead wood as characteristics worth preserving (see page 277).

Retaining such characteristics takes more than the absence of active destruction. It is precisely because of humanity's pervasive influence that even the least changed ecosystems need help surviving in the future. Białowieża's core is so small that the dynamic processes that once drove its mosaic of different micro-ecosystems probably can't

operate as they once did. Some of its large mammals are extinct. Many new species have arrived through human agency. And climate change is altering the seasonal timing and hydrological cycles of the forest.

The only alternative is proactive management — by humans. Already, conservationists in some forests set small fires to burn out underbrush before it reaches levels that could produce catastrophic fires. They shoot prey species whose populations are out of control because the top predators have been exterminated. And they have begun to control water flows into wetlands where the natural flow has been disrupted. In the future, as climate change takes hold, management may become even more radical. Some ecologists are beginning to talk about moving slowly dispersing plants and animals pole-wards or upslope to keep them in climates they can thrive in, or introducing non-native 'functional equivalents' in some ecosystems to play certain key roles.

Such talk will undoubtedly raise hackles among those ecologists for whom intervention in natural ecosystems is anathema. Yet our species' all-pervasive impact on this planet has already doomed that hands-off approach to failure.

Unfortunately, would-be managers of natural regions still know very little about how to save natural places without continuing *Homo*

sapiens' legacy of destruction. Ecologists have conventionally studied the workings of intact ecosystems, but have focused much less attention on how to keep them intact. Scientific research on the best ways to manage natural ecosystems needs to become a much higher priority.

Meanwhile, economists, ecologists and ethicists need to seek ways to bring natural ecosystems into the economic system, instead of just assuming that they exist outside of or in opposition to economics. If nothing else, this will require continued research on how to put a fair economic value on ecosystems that provide humankind with services — a classic example being wetlands that absorb storm run-off and help prevent flooding — while not dooming ecosystems such as deserts and tundras that contribute in a less obvious way.

For now, the custodians of Białowieża are letting the never-logged core area alone, even going so far as to prohibit entry to tourists except when accompanied by a guide. But the day may come when hands-off means waving goodbye. Will science know how to save Białowieża when that day comes? ■

"Scientific research on the best ways to manage natural ecosystems needs to become a much higher priority."

Virtues of visualization

Mapping techniques, used with care, can offer fresh insights into data about the world around us.

Maps are powerful, but imperfect, visualizations of the world. They are never true in any absolute sense; rather, they express certain aspects of the truth in ways that are useful for the task at hand, using rules and conventions that have to be understood in context. No flat map of Earth's curved surface can preserve both area and angle; some aspect of the reality is always distorted. So geographers have invented a plethora of projections that allow them to choose the distortions judiciously. One projection might sacrifice the accuracy of shapes and areas to preserve bearings, for example, whereas another might facilitate precision measurements in a field survey. But even this panoply often falls short — and can be downright misleading — when it comes to representing more abstract forms of information.

A classic example is the use of maps to visualize scientific and social data about nations, states or counties. Mapmakers usually do this by simply adding, say, a colour scale to represent the value of indicators such as disease outbreaks. In addition, because this technique by itself is often misleading — it takes no account of differences in population density, for example — the data are frequently normalized to per capita values. But this means that the map loses crucial data — in this case, total incidences.

The cartogram technique, in which the sizes and shapes of geographical areas are distorted to represent the population or some other variable, can overcome these issues and preserve the full richness of the data (see page 270). Used correctly, it can free statistics from the shackles of geography, allowing richer visualizations of anything from

election results, to census tabulations, to species biodiversity.

Cartograms have yet to be widely adopted, despite being the best choice for many sorts of data. One reason is that initially they can seem confusing, as most viewers' brains are more at home with the size and shape of countries and states as they appear in an atlas. Another is that efficient algorithms to create not only cartograms, but many other forms of visualizations, have generally been the preserve of a handful of specialized laboratories. That is now changing.

IBM's Visual Communication Lab in Cambridge, Massachusetts, for example, has created Many Eyes (www.many-eyes.com). This free site provides the public with tools to create visualizations such as network diagrams, which depict nodes and connections within networks, and treemaps, which display hierarchical data as groups of nested rectangles. Similarly, Google Earth and other virtual globes are providing scientists and non-scientists with unprecedented tools for geographical visualization of data. The obvious, and perhaps unexpected, enthusiasm of the public for such visualizations suggests that it is a rich vein for educators and scientists, both to explain their own work and as a means to engage young minds in critical analysis of data.

Visualizations are not a panacea. The adage 'rubbish in, rubbish out' still applies. But when used well — and, in the future, when combined with emerging surface-computing and other interactive displays — visualization can provide different views of data that force us to ask new questions, and generate fresh hypotheses.

The flood of data now coming online and the emergence of new forms — such as the data on social networks, e-mail and mobile-phone patterns that are rejuvenating the social sciences — means that visualization will be increasingly important for scientists. Such diverse windows on data should also strengthen civil society by giving scientists and citizens alike the power to sift through the data generated by governments and other institutions, and to challenge their and our own preconceptions of the world. ■

operate as they once did. Some of its large mammals are extinct. Many new species have arrived through human agency. And climate change is altering the seasonal timing and hydrological cycles of the forest.

The only alternative is proactive management — by humans. Already, conservationists in some forests set small fires to burn out underbrush before it reaches levels that could produce catastrophic fires. They shoot prey species whose populations are out of control because the top predators have been exterminated. And they have begun to control water flows into wetlands where the natural flow has been disrupted. In the future, as climate change takes hold, management may become even more radical. Some ecologists are beginning to talk about moving slowly dispersing plants and animals pole-wards or upslope to keep them in climates they can thrive in, or introducing non-native 'functional equivalents' in some ecosystems to play certain key roles.

Such talk will undoubtedly raise hackles among those ecologists for whom intervention in natural ecosystems is anathema. Yet our species' all-pervasive impact on this planet has already doomed that hands-off approach to failure.

Unfortunately, would-be managers of natural regions still know very little about how to save natural places without continuing *Homo*

sapiens' legacy of destruction. Ecologists have conventionally studied the workings of intact ecosystems, but have focused much less attention on how to keep them intact. Scientific research on the best ways to manage natural ecosystems needs to become a much higher priority.

Meanwhile, economists, ecologists and ethicists need to seek ways to bring natural ecosystems into the economic system, instead of just assuming that they exist outside of or in opposition to economics. If nothing else, this will require continued research on how to put a fair economic value on ecosystems that provide humankind with services — a classic example being wetlands that absorb storm run-off and help prevent flooding — while not dooming ecosystems such as deserts and tundras that contribute in a less obvious way.

For now, the custodians of Białowieża are letting the never-logged core area alone, even going so far as to prohibit entry to tourists except when accompanied by a guide. But the day may come when hands-off means waving goodbye. Will science know how to save Białowieża when that day comes? ■

"Scientific research on the best ways to manage natural ecosystems needs to become a much higher priority."

Virtues of visualization

Mapping techniques, used with care, can offer fresh insights into data about the world around us.

Maps are powerful, but imperfect, visualizations of the world. They are never true in any absolute sense; rather, they express certain aspects of the truth in ways that are useful for the task at hand, using rules and conventions that have to be understood in context. No flat map of Earth's curved surface can preserve both area and angle; some aspect of the reality is always distorted. So geographers have invented a plethora of projections that allow them to choose the distortions judiciously. One projection might sacrifice the accuracy of shapes and areas to preserve bearings, for example, whereas another might facilitate precision measurements in a field survey. But even this panoply often falls short — and can be downright misleading — when it comes to representing more abstract forms of information.

A classic example is the use of maps to visualize scientific and social data about nations, states or counties. Mapmakers usually do this by simply adding, say, a colour scale to represent the value of indicators such as disease outbreaks. In addition, because this technique by itself is often misleading — it takes no account of differences in population density, for example — the data are frequently normalized to per capita values. But this means that the map loses crucial data — in this case, total incidences.

The cartogram technique, in which the sizes and shapes of geographical areas are distorted to represent the population or some other variable, can overcome these issues and preserve the full richness of the data (see page 270). Used correctly, it can free statistics from the shackles of geography, allowing richer visualizations of anything from

election results, to census tabulations, to species biodiversity.

Cartograms have yet to be widely adopted, despite being the best choice for many sorts of data. One reason is that initially they can seem confusing, as most viewers' brains are more at home with the size and shape of countries and states as they appear in an atlas. Another is that efficient algorithms to create not only cartograms, but many other forms of visualizations, have generally been the preserve of a handful of specialized laboratories. That is now changing.

IBM's Visual Communication Lab in Cambridge, Massachusetts, for example, has created Many Eyes (www.many-eyes.com). This free site provides the public with tools to create visualizations such as network diagrams, which depict nodes and connections within networks, and treemaps, which display hierarchical data as groups of nested rectangles. Similarly, Google Earth and other virtual globes are providing scientists and non-scientists with unprecedented tools for geographical visualization of data. The obvious, and perhaps unexpected, enthusiasm of the public for such visualizations suggests that it is a rich vein for educators and scientists, both to explain their own work and as a means to engage young minds in critical analysis of data.

Visualizations are not a panacea. The adage 'rubbish in, rubbish out' still applies. But when used well — and, in the future, when combined with emerging surface-computing and other interactive displays — visualization can provide different views of data that force us to ask new questions, and generate fresh hypotheses.

The flood of data now coming online and the emergence of new forms — such as the data on social networks, e-mail and mobile-phone patterns that are rejuvenating the social sciences — means that visualization will be increasingly important for scientists. Such diverse windows on data should also strengthen civil society by giving scientists and citizens alike the power to sift through the data generated by governments and other institutions, and to challenge their and our own preconceptions of the world. ■

RESEARCH HIGHLIGHTS

**Active aphids**

Proc. R. Soc. Lond. B doi: 10.1098/rspb.2008.0880 (2008)

Airborne arthropods are often divided into active fliers, which travel under their own power, and passive drifters on air currents. Now Andy Reynolds of Rothamsted Research in Harpenden, UK, and a colleague announce that aphids and other small-winged insects should not be labelled passive, because they can control their altitude.

In simulations of air currents and aphid behaviour, results were closest to the observed distribution of aphids when the insects flew so as to magnify the effect of air movements. In updrafts, the authors claim, they fly just enough to counterbalance gravity. In downdrafts, they stop flapping and sink.

ROTHAMSTED RESEARCH

BIOCHEMISTRY**Protecting the heart**

Science **321**, 1493–1495 (2008)

A protein that reduces heart damage in rats could hold the key to protecting humans against a potentially life-threatening condition: reduction of blood flow to the heart, or cardiac ischaemia.

Daria Mochly-Rosen, of Stanford University in California, and her colleagues found that in rats, increased activity of a protein called aldehyde dehydrogenase 2 (ALDH2) is correlated with reduced damage following cardiac ischaemia. The researchers then isolated a compound called Alda-1 that activates ALDH2, also reducing heart damage by ischaemia.

Two-fifths of East Asians carry a mutation in the *ALDH2* gene that leads to reduced ALDH2 activity. Alda-1 activated the mutant protein, and restored it to normal levels of activity.

ASTRONOMY**Fossils on Titan**

Science **321**, 1475–1478 (2008)

Before 13 June 2007, astronomers observed Saturn's moon Titan (pictured right) only when it was inside the planet's magnetic field. But on that day the probe Cassini observed Titan outside the controlling influence of Saturn's magnetosphere, in a region permeated by the Sun's magnetic field.

Cesar Bertucci, now at the Institute for Astronomy and Space Physics in Buenos Aires, and his colleagues report that Titan, which lacks a major magnetic field of its own,

retains an imprint of Saturn's magnetic fields after leaving the planet's magnetosphere. This occurs in the form of 'fossil' fields frozen into the plasma surrounding the moon. The solar magnetic field then overlaid these fossil fields, which persisted for between 20 minutes and three hours, they report.

CONSERVATION**Homes on the range**

Conserv. Biol. **22**, 912–921 (2008)

When it comes to choosing where to live, the best environmental intentions can lead to the worst results.

M. Nils Peterson, now at North Carolina State University in Raleigh, and his colleagues surveyed more than 400 households in Idaho and Wyoming. They found that the least educated people who cared least about the environment chose to live in established residential areas, with low environmental

impacts. But people with high levels of education and environmental concerns chose homes in environmentally sensitive natural areas. Income was not a factor. Strangely, the longer residents lived in a natural area, the less they cared about the environment.

EVOLUTION**Sexy and doomed**

Proc. Natl Acad. Sci. USA doi: 10.1073/

pnas.0803851105 (2008)

Why haven't oncogenes — genes that cause cancer — been eliminated by natural selection? André Fernandez and a colleague from Ohio University in Athens believe they have found a possible answer. In a species of swordtail fish, *Xiphophorus cortezi*, the oncogene *Xmrk* also causes a dark pattern on male fish's tail fin that females in two out of three populations studied found attractive. Indeed, fish that have developed cancer had the largest patterns and were able to mate for many months before the disease killed them.

NEUROLOGY**Dopamine link to fragile X**

Neuron doi: 10.1016/j.neuron.2008.06.027 (2008)

The lack of fragile X mental retardation protein (FMRP) causes a disorder with physical, cognitive and behavioural symptoms. FMRP has many roles, and Min Zhuo at the University of Toronto and his colleagues now add another: relaying messages in the dopamine pathway that shapes memory, planning and attention.

Zhuo and his colleagues studied cultured neurons and mice lacking FMRP. The

NASA-JPL

RESEARCH HIGHLIGHTS

**Active aphids**

Proc. R. Soc. Lond. B doi: 10.1098/rspb.2008.0880 (2008)

Airborne arthropods are often divided into active fliers, which travel under their own power, and passive drifters on air currents. Now Andy Reynolds of Rothamsted Research in Harpenden, UK, and a colleague announce that aphids and other small-winged insects should not be labelled passive, because they can control their altitude.

In simulations of air currents and aphid behaviour, results were closest to the observed distribution of aphids when the insects flew so as to magnify the effect of air movements. In updrafts, the authors claim, they fly just enough to counterbalance gravity. In downdrafts, they stop flapping and sink.

ROTHAMSTED RESEARCH

BIOCHEMISTRY**Protecting the heart**

Science **321**, 1493–1495 (2008)

A protein that reduces heart damage in rats could hold the key to protecting humans against a potentially life-threatening condition: reduction of blood flow to the heart, or cardiac ischaemia.

Daria Mochly-Rosen, of Stanford University in California, and her colleagues found that in rats, increased activity of a protein called aldehyde dehydrogenase 2 (ALDH2) is correlated with reduced damage following cardiac ischaemia. The researchers then isolated a compound called Alda-1 that activates ALDH2, also reducing heart damage by ischaemia.

Two-fifths of East Asians carry a mutation in the *ALDH2* gene that leads to reduced ALDH2 activity. Alda-1 activated the mutant protein, and restored it to normal levels of activity.

ASTRONOMY**Fossils on Titan**

Science **321**, 1475–1478 (2008)

Before 13 June 2007, astronomers observed Saturn's moon Titan (pictured right) only when it was inside the planet's magnetic field. But on that day the probe Cassini observed Titan outside the controlling influence of Saturn's magnetosphere, in a region permeated by the Sun's magnetic field.

Cesar Bertucci, now at the Institute for Astronomy and Space Physics in Buenos Aires, and his colleagues report that Titan, which lacks a major magnetic field of its own,

retains an imprint of Saturn's magnetic fields after leaving the planet's magnetosphere. This occurs in the form of 'fossil' fields frozen into the plasma surrounding the moon. The solar magnetic field then overlaid these fossil fields, which persisted for between 20 minutes and three hours, they report.

CONSERVATION**Homes on the range**

Conserv. Biol. **22**, 912–921 (2008)

When it comes to choosing where to live, the best environmental intentions can lead to the worst results.

M. Nils Peterson, now at North Carolina State University in Raleigh, and his colleagues surveyed more than 400 households in Idaho and Wyoming. They found that the least educated people who cared least about the environment chose to live in established residential areas, with low environmental

impacts. But people with high levels of education and environmental concerns chose homes in environmentally sensitive natural areas. Income was not a factor. Strangely, the longer residents lived in a natural area, the less they cared about the environment.

EVOLUTION**Sexy and doomed**

Proc. Natl Acad. Sci. USA doi: 10.1073/

pnas.0803851105 (2008)

Why haven't oncogenes — genes that cause cancer — been eliminated by natural selection? André Fernandez and a colleague from Ohio University in Athens believe they have found a possible answer. In a species of swordtail fish, *Xiphophorus cortezi*, the oncogene *Xmrk* also causes a dark pattern on male fish's tail fin that females in two out of three populations studied found attractive. Indeed, fish that have developed cancer had the largest patterns and were able to mate for many months before the disease killed them.

NEUROLOGY**Dopamine link to fragile X**

Neuron doi: 10.1016/j.neuron.2008.06.027 (2008)

The lack of fragile X mental retardation protein (FMRP) causes a disorder with physical, cognitive and behavioural symptoms. FMRP has many roles, and Min Zhuo at the University of Toronto and his colleagues now add another: relaying messages in the dopamine pathway that shapes memory, planning and attention.

Zhuo and his colleagues studied cultured neurons and mice lacking FMRP. The

NASA-JPL

neurons responded sluggishly to stimulation of a dopamine receptor known as D1. The mice had an altered distribution pattern of the G protein-coupled receptor kinase 2 protein, GRK2, which decreases D1's activity. FMRP and GRK2 were also found to bind to one another, and blocking GRK2 with a drug corrected the mice's D1 signalling deficiency. Drugs that boosted D1 signalling also restored some behaviour deficits in the mice. Zhuo and his colleagues conclude that FMRP is a player in the dopamine signalling architecture of the forebrain and that this may explain some of the problems caused by its deficit in fragile X syndrome.

CHEMISTRY

Ro-taxing synthesis

Angew. Chem. Int. Edn doi: 10.1002/anie.200803056 (2008)

Chemists are fascinated by rotaxanes — a class of compound made from a dumbbell-shaped chemical inside a molecular ring — because such structures could be useful in making microscale machines. Sheng-Hsien Chiu and his colleagues at the National Taiwan University have now created the smallest rotaxane to date (pictured right), made from just 76 atoms, with a molecular mass of 510 daltons.

Their ring is a crown ether — a circle of alternating oxygen atoms and ethylene groups — and their dumbbell's axle is a dialkylammonium ion. After threading the axle through the ring, grinding the mixture up with a tetrazine caps the axle with pyridazine groups to make the rotaxane.

The scientists hope that they can shave a few atoms off the rotaxane to beat their own record.

EVOLUTIONARY BIOLOGY

Eggs in many baskets

Proc. R. Soc. Lond. B doi: 10.1098/rspb.2008.0794 (2008)

In Bibron's toadlet (*Pseudophryne bibronii*) males are the nestmakers and females shop around. Females of this Australian amphibian lay their eggs in the nests of up to eight males. This makes it the first frog or toad in which such 'sequential polyandry'

has been genetically proven, and the most extreme example of this mating tactic seen in any vertebrate.

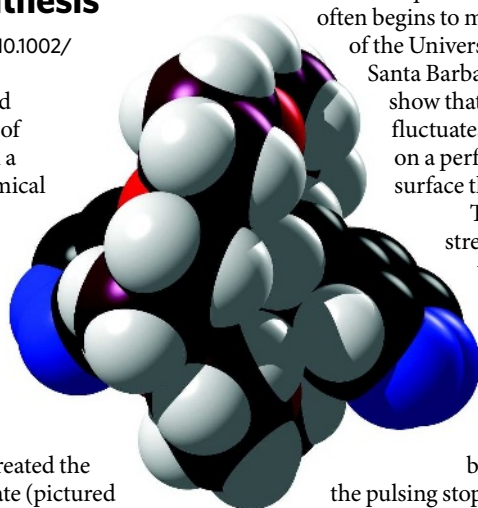
Using DNA analysis and observations, Phillip Byrne, of Monash University in Melbourne, and his colleague showed that females that laid eggs in more nests had more surviving tadpoles. They suggest that polyandry is an insurance policy against nests flooding or drying out, and that other egg-layers in which dad tends the brood in a risky environment may do the same.

FLUID DYNAMICS

Making water wander

Phys. Rev. Lett. **101**, 114501 (2008)

As a stream of liquid flows down a slope, it often begins to meander. Björn Birnir of the University of California, Santa Barbara, and his colleagues show that if the liquid flow fluctuates this can happen even on a perfectly smooth, clean surface that does not erode.



The researchers ran a stream of a water mixed with glycerine down a plastic sheet. As the flow changed from smooth to pulsed, the rivulet began meandering. The wavy path can be sustained once the pulsing stops, as the stream hits droplets left on the surface during its earlier wanderings.

ATMOSPHERIC CHEMISTRY

Glasses in the sky

Atmos. Chem. Phys. **8**, 5221–5244 (2008)

Some solutions of water-soluble organic compounds form glasses below -43°C . The upper troposphere, where cirrus clouds form, is this cold; glass formation here could impede the water uptake of aerosol particles, ice nucleation and crystal growth, with implications for cloud formation.

Thomas Koop of Bielefeld University in Germany and his colleagues investigated glass formation in various solutions under realistic atmospheric conditions in the lab. They found that aerosol particles enriched with large organic molecules are most likely to form glasses at low temperatures and high relative humidity. Modelling studies should be used to assess the regional and global impact of atmospheric glass formation, say the authors.

JOURNAL CLUB

Nicola Hamilton and David Attwell
University College London

Two neuroscientists are surprised by the link between a brain-chemical transporter and sexual orientation.

Many nerve cells in the brain release the chemical neurotransmitter glutamate to signal to other neurons via receptors. Dedicated transporters then remove glutamate from the extracellular space to end signalling.

Cystine-glutamate exchangers are unusual glutamate transporters because they do the reverse, adding glutamate to the extracellular space while removing cystine. David Featherstone of the University of Illinois, Chicago, and his colleagues have found that in the fruitfly *Drosophila melanogaster*, knocking down expression of a cystine-glutamate exchanger in non-neuronal glial cells leads to a dramatic change in the sexual behaviour of male flies: they mate with both males and females owing to altered processing of sex-specific chemosensory cues (Y. Grosjean *et al. Nature Neurosci.* **11**, 54–61; 2008).

This behaviour may be caused by an increase in the number of glutamate signalling receptors, which is induced by the fall in extracellular glutamate concentration that follows transporter knockdown. Indeed, the effect of the knockdown could be reversed by feeding the flies a drug that reduces glutamate signalling, and could be mimicked by feeding normal flies a drug that enhances glutamate signalling.

These studies raise questions about whether human sexual orientation, long assumed to be due to a mix of genes and environment, could also be altered by perturbations of neurotransmitter signalling. Could differences in such signalling contribute to different sexual preferences?

The possibility of altering sexual preference pharmacologically is worrying. We cannot rule out a future regression to the twentieth-century idea that sexual behaviour should be regulated by society.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS

Green issues dominate election

VANCOUVER

As the United States heads for a showdown at the polls, its Canadian neighbour is rushing out its own election — announced on 7 September, voting is scheduled for 14 October. In Canada, a country dominated by vast expanses of wilderness — as well as growing greenhouse-gas emissions — green issues are at the forefront of the debate.

"In many ways this is the first time that this has happened in Canadian history — the first environment election and the first climate election," says Graham Saul, executive director of the non-partisan Climate Action Network in Ottawa.

Prime Minister Stephen Harper, the leader of Canada's minority Conservative (centre-right) government, has so far campaigned on his ability to manage the economy. Opposing Liberal (centre-left) party leader Stéphane Dion's defining policy is the 'Green Shift', a plan to shift taxes away from income and investment and towards pollution and greenhouse-gas emissions. Dion has attacked the Conservative party's environmental inaction, including its abandonment of Kyoto Protocol targets in 2006. He even has a dog named Kyoto, which

features in the party's television advertisements. A quote from Al Gore — "We should tax what we burn, not what we earn" — is prominent on the Green Shift website.

Harper was elected in 2006 after a decade of Liberal leadership. The most recent polls show that the Conservatives are ahead in this year's race, in which they hope to win a majority government.

The focus on green issues is exemplified by a row over the Green party's participation in televised debates scheduled for the first week of October. Despite having a member of parliament and more than 5% support in polls (the usual criteria for

participating in the event), when Green leader Elizabeth May asked for her invitation, Harper threatened to boycott the event if she attended. "Ms May is basically presenting the same positions as Mr Dion," complained Harper, saying that would make the event unfair. This ignited a storm of controversy and drew pointed remarks about Harper's apparent unwillingness to be questioned on his environmental policies. Under pressure, Harper reversed his position and last week May was invited onto the show.

It was a busy week. On 9 September, four former prime ministers — two Conservative and two Liberal — signed the Canadians for Climate Leadership Declaration. The document is a general call for action without many specifics, but it does make a key point that pricing carbon — they support Can\$30 (US\$28) per tonne CO₂ equivalent — is an essential part of tackling climate change.

High emissions

Canada was one of the first countries to sign the Kyoto Protocol in 1998, pledging to reduce emissions by 6% below 1990 baseline levels by 2012. But emissions were more than 20% over baseline when the Conservatives took power in 2006, just before they admitted that Kyoto targets wouldn't be met. Today Canada's per capita emissions and energy consumption are among the highest in the world. Rapid population growth and rising export markets make national emission targets — rather than per capita targets — hard to meet, notes economist and policy expert Thomas Courchene of Queen's University in Kingston, Ontario. The country also has strong commercial interests in fishing, forestry, trucking and oil, including Alberta's tar-sand oil refineries, which can make environmental policies hard to push through.



F. CHARTRAND/THE CANADIAN PRESS/AP PHOTO

Conservative leader Stephen Harper puts the economy high on his agenda.

The Green Shift's main promise is to tax greenhouse-gas emissions, starting at Can\$10 per tonne of CO₂ and reaching Can\$40 per tonne within four years. This would be matched by reductions in income tax, making the plan revenue neutral. Existing taxes on gasoline mean that the plan shouldn't hit pump prices, and there will be no price hikes for diesel and aviation fuel for the first year. "It will mean huge price increases for coal, and increased interest in clean coal," says Courchene. Harper has promised to cut diesel taxes.

The Conservatives don't dispute that climate change is important, but argue over the way to tackle it. Their policies include cutting emissions to 20% below 2006 levels by 2020, requiring carbon capture and storage for Alberta tar-sand operations that start after 2012, and starting an emissions-trading plan based on national energy efficiency, using energy intensity — the amount of energy per unit of gross domestic product — rather than absolute caps.

The Liberal party aims for a 20% emissions cut below 1990 levels by 2020. And it has plans for a cap-and-trade programme, as exists in Europe, but acknowledges it will take years to set up the bureaucracy to support it; a tax can be implemented immediately. The main point emphasized by the declaration of former prime ministers, advocacy groups and all the main Canadian parties aside from the Conservatives,



Liberal leader Stéphane Dion's election campaign focuses on green issues.

J. WOODS/THE CANADIAN PRESS/AP PHOTO



HAVE YOUR SAY
Comment on any of our
News stories, online.
www.nature.com/news

says Saul, is to put a price on carbon quickly.

A small foray into carbon taxing has already been made in Canada. British Columbia's Liberal provincial government implemented the first major North American carbon tax in July, at a rate of Can\$10 per tonne of CO₂ (or equivalent emissions), rising to Can\$30 per tonne in 2012. The province's plan is also revenue neutral, with companies paying the majority of the tax and individuals receiving most of the credit — including a one-off cheque for Can\$100 earlier this summer that is meant to be spent on creating more environmentally friendly lifestyles.

Earlier this month, more than 60% of 1,000 people polled nationally said that climate change would be important or the most important issue in deciding their vote. Such opinions don't always survive in the polling booth, but constant reminders of the impacts of climate change did influence Conservative-minded Australians to vote out their government in elections last year.

Canadian policy experts see their country lagging behind Europe on climate issues. It could yet catch up. But although Canada's election precedes the one down south, their policies may yet be held back by US decisions. "There's a real danger that we're just going to kind of adopt whatever happens in the United States," says Saul. "Canadian politicians don't like to make a stand if Americans aren't doing the same."

Nicola Jones

See Editorial, page 263.

Iran holds AIDS doctors

Iran is under mounting pressure to reveal the whereabouts of two of the nation's HIV researchers, who have been detained without charge since late June.

The brothers, Arash and Kamiar Alaei, have achieved international acclaim for their progressive HIV-prevention programme and were scheduled to speak about it at the 17th International AIDS Conference in Mexico City last month. Both were arrested before leaving the country. The men had collaborated with other scientists around the world, including some in the United States, and are not thought to have been politically active.

Protests in their support were made at the conference. Since then, several human-rights organizations, including Physicians for Human Rights and Amnesty International, have called

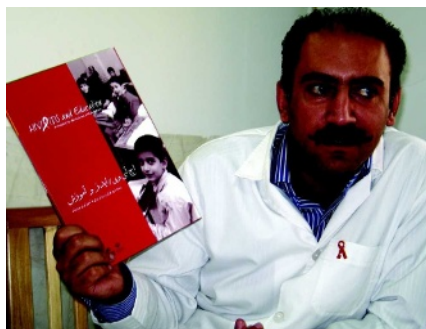
on Iran to abide by its international legal obligations to explain the arrests and allow the men access to lawyers and the right to contest their detention before a judge.

The call has been taken up by several scientific bodies, including the International AIDS Society, the Foundation for AIDS Research and the American Association for the Advancement of Science (AAAS), and thousands of scientists and physicians have signed an online petition at www.iranfreethedocs.org. Barry Bloom, dean of the Harvard School of Public Health in Boston, Massachusetts, has also expressed his "deep concern" over the brothers' detention.

Kamiar, the younger of the brothers, holds a master's degree from the Harvard School of Public Health and was to have resumed doctoral studies at the University of Albany's School of Public Health in New York. Arash, former head of international education and research cooperation at the Iranian National Research Institute of Tuberculosis and Lung Disease, runs a clinic in Tehran.

HIV is a taboo subject in much of the Middle East, including Iran, but the brothers have helped to forge impressive prevention and treatment programmes — the country is now one of the few to distribute condoms and syringes in prisons, for example.

Declan Butler



Arash Alaei has not been seen since June.

H. ALLAM/KRT/NEWS.COM

Japan fast-tracks stem-cell patent

Kyoto University in Japan has acquired the world's first patent for induced pluripotent stem (iPS) cells. The university paid registration fees for the patent (2008-131577), which is of only limited coverage and applies only in Japan, on 2 September, after it was approved by the Japanese patent office.

Shinya Yamanaka, a stem-cell researcher at the university, created the first iPS cells in 2006 by using four genetic factors to 'reprogram' adult cells from mice into an embryonic-stem-cell-like state. A year later he achieved the same in human cells, publishing the research the same day as a US team led by James Thomson of the University of Wisconsin. Thomson's group used a different set of genetic factors to make its human iPS cells.

iPS cells have the potential to develop into almost any of the body's cell types, and hold great promise in therapy and disease modelling. Kyoto University applied for an international patent (PCT/JP2006/324881) in December 2006. A US patent application is in Yamanaka's name.

Kyoto University expects the international patent to cover iPS cells from all species — "any cell, besides a germ cell, of any type of organism," says Hideya Hayashi of the university's iPS Cell Research and Application office.

The international patent has many claims in it, so it will probably take another year or longer for the various patent offices to process and approve it, says Hayashi. Kyoto University did

not want to wait, so it fast-tracked the Japanese patent application for the most basic claim, covering the "method" of using Yamanaka's four genetic factors to reprogram a cell. "A lot of people were wondering what was happening with these patents. We wanted to make it clear," Hayashi says.

Other claims in the international patent still need to be worked out, especially that covering the product, the iPS cells themselves. It's not certain whether Yamanaka's patents cover iPS cells created by reprogramming with different genetic factors. Itsuki Shimbo, a patent attorney with the Tokyo-based firm Tsubame says it is unclear whether the patent will be able to restrict scientists using other combinations.

Hayashi says the university does not plan to restrict others from doing iPS cell research. Rather, it wants to ensure that people, including Yamanaka, are not restricted by other patents and can freely use iPS cell technology. "We want to remove any potential obstacles to the quick clinical application of iPS technology," Hayashi says. "We are not trying to confine its use."

Shimbo says that Yamanaka's patent might prevent future restrictive iPS patents based on different combinations of genetic factors. "It will be very hard to claim that such a change was truly novel," he says.

The Japanese patent will be valid until 6 December 2026, 20 years from the application date.

David Cyranoski



HAVE YOUR SAY
Comment on any of our
News stories, online.
www.nature.com/news

says Saul, is to put a price on carbon quickly.

A small foray into carbon taxing has already been made in Canada. British Columbia's Liberal provincial government implemented the first major North American carbon tax in July, at a rate of Can\$10 per tonne of CO₂ (or equivalent emissions), rising to Can\$30 per tonne in 2012. The province's plan is also revenue neutral, with companies paying the majority of the tax and individuals receiving most of the credit — including a one-off cheque for Can\$100 earlier this summer that is meant to be spent on creating more environmentally friendly lifestyles.

Earlier this month, more than 60% of 1,000 people polled nationally said that climate change would be important or the most important issue in deciding their vote. Such opinions don't always survive in the polling booth, but constant reminders of the impacts of climate change did influence Conservative-minded Australians to vote out their government in elections last year.

Canadian policy experts see their country lagging behind Europe on climate issues. It could yet catch up. But although Canada's election precedes the one down south, their policies may yet be held back by US decisions. "There's a real danger that we're just going to kind of adopt whatever happens in the United States," says Saul. "Canadian politicians don't like to make a stand if Americans aren't doing the same."

Nicola Jones

See Editorial, page 263.

Iran holds AIDS doctors

Iran is under mounting pressure to reveal the whereabouts of two of the nation's HIV researchers, who have been detained without charge since late June.

The brothers, Arash and Kamiar Alaei, have achieved international acclaim for their progressive HIV-prevention programme and were scheduled to speak about it at the 17th International AIDS Conference in Mexico City last month. Both were arrested before leaving the country. The men had collaborated with other scientists around the world, including some in the United States, and are not thought to have been politically active.

Protests in their support were made at the conference. Since then, several human-rights organizations, including Physicians for Human Rights and Amnesty International, have called

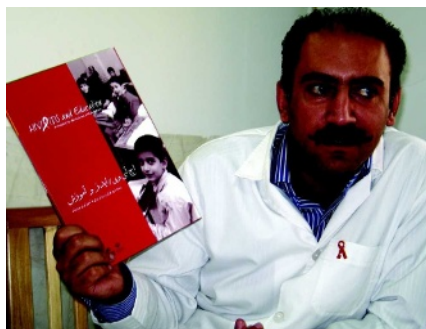
on Iran to abide by its international legal obligations to explain the arrests and allow the men access to lawyers and the right to contest their detention before a judge.

The call has been taken up by several scientific bodies, including the International AIDS Society, the Foundation for AIDS Research and the American Association for the Advancement of Science (AAAS), and thousands of scientists and physicians have signed an online petition at www.iranfreethedocs.org. Barry Bloom, dean of the Harvard School of Public Health in Boston, Massachusetts, has also expressed his "deep concern" over the brothers' detention.

Kamiar, the younger of the brothers, holds a master's degree from the Harvard School of Public Health and was to have resumed doctoral studies at the University of Albany's School of Public Health in New York. Arash, former head of international education and research cooperation at the Iranian National Research Institute of Tuberculosis and Lung Disease, runs a clinic in Tehran.

HIV is a taboo subject in much of the Middle East, including Iran, but the brothers have helped to forge impressive prevention and treatment programmes — the country is now one of the few to distribute condoms and syringes in prisons, for example.

Declan Butler



Arash Alaei has not been seen since June.

H. ALLAN/KRT/NEWS.COM

Japan fast-tracks stem-cell patent

Kyoto University in Japan has acquired the world's first patent for induced pluripotent stem (iPS) cells. The university paid registration fees for the patent (2008-131577), which is of only limited coverage and applies only in Japan, on 2 September, after it was approved by the Japanese patent office.

Shinya Yamanaka, a stem-cell researcher at the university, created the first iPS cells in 2006 by using four genetic factors to 'reprogram' adult cells from mice into an embryonic-stem-cell-like state. A year later he achieved the same in human cells, publishing the research the same day as a US team led by James Thomson of the University of Wisconsin. Thomson's group used a different set of genetic factors to make its human iPS cells.

iPS cells have the potential to develop into almost any of the body's cell types, and hold great promise in therapy and disease modelling. Kyoto University applied for an international patent (PCT/JP2006/324881) in December 2006. A US patent application is in Yamanaka's name.

Kyoto University expects the international patent to cover iPS cells from all species — "any cell, besides a germ cell, of any type of organism," says Hideya Hayashi of the university's iPS Cell Research and Application office.

The international patent has many claims in it, so it will probably take another year or longer for the various patent offices to process and approve it, says Hayashi. Kyoto University did

not want to wait, so it fast-tracked the Japanese patent application for the most basic claim, covering the "method" of using Yamanaka's four genetic factors to reprogram a cell. "A lot of people were wondering what was happening with these patents. We wanted to make it clear," Hayashi says.

Other claims in the international patent still need to be worked out, especially that covering the product, the iPS cells themselves. It's not certain whether Yamanaka's patents cover iPS cells created by reprogramming with different genetic factors. Itsuki Shimbo, a patent attorney with the Tokyo-based firm Tsubame says it is unclear whether the patent will be able to restrict scientists using other combinations.

Hayashi says the university does not plan to restrict others from doing iPS cell research. Rather, it wants to ensure that people, including Yamanaka, are not restricted by other patents and can freely use iPS cell technology. "We want to remove any potential obstacles to the quick clinical application of iPS technology," Hayashi says. "We are not trying to confine its use."

Shimbo says that Yamanaka's patent might prevent future restrictive iPS patents based on different combinations of genetic factors. "It will be very hard to claim that such a change was truly novel," he says.

The Japanese patent will be valid until 6 December 2026, 20 years from the application date.

David Cyranoski

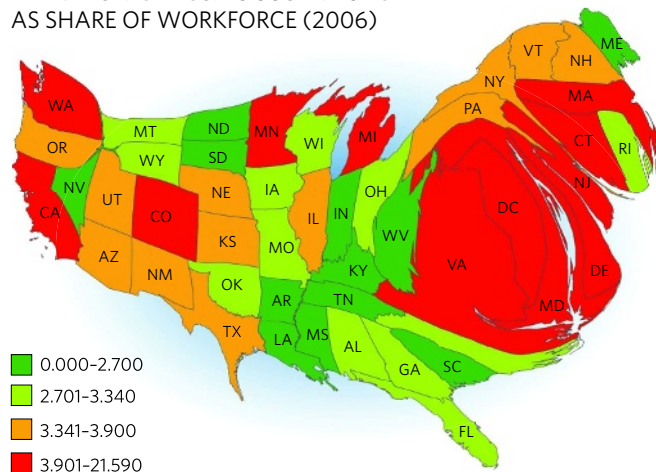


BIRDS CONFOUND CLIMATE MODELS

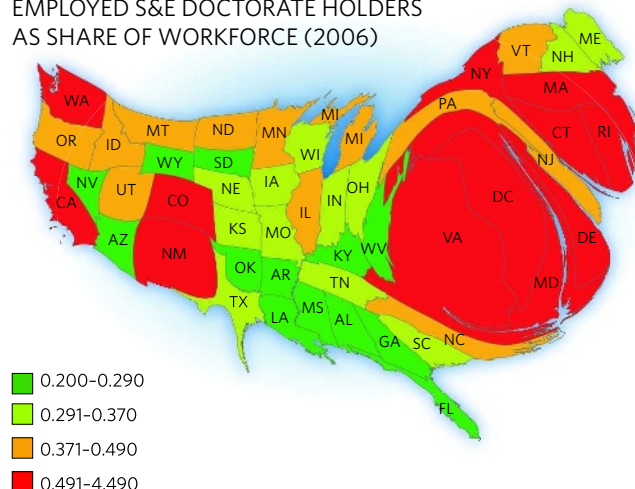
Could more than 60% of 'climate envelope' studies be wrong?

www.nature.com/news

INDIVIDUALS IN S&E OCCUPATIONS
AS SHARE OF WORKFORCE (2006)



EMPLOYED S&E DOCTORATE HOLDERS
AS SHARE OF WORKFORCE (2006)



Science and engineering jobs have the biggest share of the workforce in the District of Columbia and the adjacent states of Maryland and Virginia.

This indicator shows a state's ability to attract and retain highly trained scientists and engineers.

for generating cartograms, which was the one used here. The software implementation is the open-source ScapeToad, released in May by the Chôros Laboratory at the Swiss Federal Institute of Technology (EPFL) in Lausanne, which researches the concept of space in society, from urban planning to territorial development.

Cartograms have rarely been used for science indicators, although the NSF generated two for its 2006 science and engineering indicators. "The resulting feedback was mixed," says John Gawalt of the NSF's statistics division, "they can be difficult for readers to interpret."

But visualizing information in this way can be stimulating. "For such a familiar map as the United States, the combined distortions and

colour scales of these cartograms are highly meaningful; the maps become somehow shocking and subversive," says Rémi Barré, a researcher at the National Conservatory of Arts and Crafts (CNAM) in Paris and former head of the French science indicators agency, the Paris-based Science and Technology Observatory (OCT). "This double visualization is powerful and welcome."

"They're a really nice visualization of some things I hadn't really appreciated before," says Newman. "The one of R&D as share of gross domestic product is pretty close to a population cartogram, except for Colorado and New Mexico, where the big national labs are. It's interesting to see Michigan, with the auto industry, and

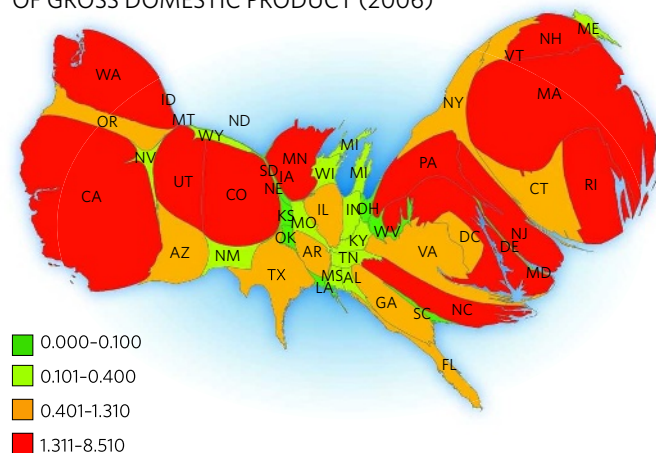
Washington state, home of Microsoft, looming large on the industry-funding cartogram. Why is Oregon so big? Is that Intel?"

Geographer Jacques Lévy, director of the Chôros lab, says the *Nature* maps are fascinating. "I'm struck by the efficient resistance of the east, including a part of the Rust Belt and the coastal south. Conversely, California and the west coast do not seem to do as well as we might imagine."

"The maps also show how in a supposed neo-liberal America, the role of a powerful, multi-tier state can challenge the spontaneous distribution of resources induced by socio-economic rationales," Lévy adds. ■

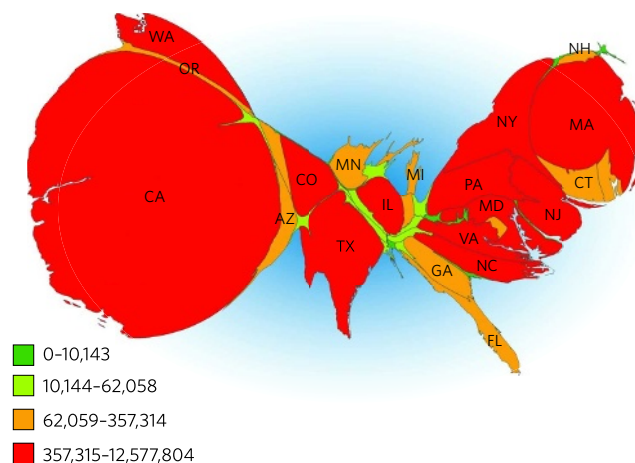
See Editorial, page 264.

VENTURE CAPITAL DISBURSED PER \$1,000
OF GROSS DOMESTIC PRODUCT (2006)



High values of venture capital normalized to the size of a state's economy show which states are successfully attracting venture capital to fuel growth.

VENTURE CAPITAL DISBURSED (\$ THOUSANDS) (2006)



Absolute values of venture capital disbursed show that it is concentrated in a handful of states, with California obtaining almost half.

Q&A



INNOVATIVE IDEAS

GARRISON PHOTOGRAPHY

Data on display

Risking being scooped and having patents refused, some scientists are posting their data online as they produce them. Organic chemist **Jean-Claude Bradley** (left) of Drexel University in Philadelphia, Pennsylvania, and biochemist **Cameron Neylon** of the University of Southampton, UK, describe this 'open notebook' approach.



DREXEL UNIV.; C. NEYLON

What is an 'open notebook'?

Bradley: The basic philosophy of open-notebook science is to have no insider information. Essentially all the information that is available to the [research] group is available to the rest of the world. You have an objective, a procedure and a log section, in which you report what you actually do. Then you've got your raw data files, and you link to those. That is all the information that any scientist needs to be able to figure out what you did and to analyse it. I use a wiki space that gives me a time stamp on the entries.

Neylon: We're aiming to improve scientific communication. The ultimate form would be if everything were available as it happened. That's never going to be for everybody. You still have issues such as patient privacy, the safety of people doing animal experimentation and so on. So in some cases, data shouldn't be made fully or immediately available.

What are the main concerns?

Neylon: The main issue is the fear of rivals stealing data. The second one is: will I be able to publish? And that depends on the publisher. Most publishers regard what we do as the equivalent of presenting at a conference, or a preprint. That hasn't been tested across a wide range of publishers, and there's at least one — the American Chemical Society — that doesn't allow prepublication in any form whatsoever. There's also a legitimate concern that a lot of people will put out a lot of rubbish.

Is this going to make traditional publishing obsolete?

Bradley: No. I'm publishing a paper [based on work that is openly available on my online notebook]. The notebook is about publishing data as quickly as possible. The paper is about synthesizing knowledge from all those results. But we want the best of both worlds, so we want to publish using traditional channels and we want to link back to the notebook.

Are scientists working in industry interested in the idea?

Neylon: The people within companies who are trying to do this are finding it a very hard sell to the board but it is being talked about.

Bradley: If patenting is important to you, you cannot do open-notebook science, it's that simple. But in big pharma, there's an awful lot of background data, which they call precompetitive work, which are really valuable to other people but end up not being as valuable to the company.

Has your open notebook ever been used to claim priority over a discovery?

Neylon: On our blog we had a statement with evidence that a piece of research worked before a paper from another group came out, although probably after the paper was submitted. I'm interested in putting the statement "we were the first to report it" in our paper when we publish the work and seeing what the response is. People's views differ about whether that would be reasonable.

When will the idea become more popular?

Neylon: Open notebooks are practical but tough at the moment. My feeling is that the tools are not yet easy enough to use. But I would say that a larger proportion of people will be publishing electronically and openly in ten years.

Bradley: There are different ways to do this. If you use a wiki-based approach, that's something you could do overnight. But the whole lab needs to realize that it's something they need to do.

Neylon: An important point that sometimes get missed about electronic systems is that other people might be looking at them, so they tend to force much higher standards of record-keeping. The record becomes much better and more flexible, but it involves a lot more work to keep it up to date.

Interview by Katherine Sanderson

In the third of our election-themed podcasts available online, *Nature* looks at where US innovation policy might go under a new president. Excerpts from our panel discussion:

"One of the things I would like to see the candidates understand — I think Obama gets it intuitively, and McCain may be coming around to it — is that technology and innovation drive long-term economic growth... Real national leadership on these issues, on a consistent and ongoing basis, would be what I want to see from the next administration."

Stephen Ezell, senior analyst, Information Technology and Innovation Foundation, Washington DC

"It's astounding that over the past eight years, the number of computer-science graduates in the United States has decreased by 50%, at a time when computer science became one of the most high-growth job industries."

Stephen Ezell

"It is truly a competition for the best and brightest right now, and we need to do that. If it means stapling a green card to the diploma for people who have a degree in science that we think is valuable, then we should be looking to do it now."

Bill Bates, vice-president, government affairs, Council on Competitiveness, Washington DC

"We need to find ways in this country to get over being so worried about supporting industrial policy and funding, that Valley of Death that exists between our universities and the venture capitalists."

Tobin Smith, associate vice-president for federal relations, Association of American Universities, Washington DC

To hear the full discussion, chaired by our columnist David Goldston, visit www.nature.com/nature/podcast.



A longer paper gathers more citations

Researchers could garner more citations simply by making their papers longer, a study seems to imply.

In an analysis of 30,027 peer-reviewed papers published between 2000 and 2004 in top astronomy journals, astronomer Krzysztof Stanek of Ohio State University in Columbus found that the median number of citations increases with the length of the paper — from just 6 for papers of 2–3 pages to about 50 for 50-page papers¹.

There is, however, a limit to the benefits of size: citations start to tail off when papers reach lengths of 80 pages or so, perhaps because fewer people have the stamina to read them.

It is unexpected, says astronomer Jörg Dietrich of the European Southern Observatory

headquarters in Germany, who recently conducted a similar analysis and found the same results but didn't publish them. "I expected that shorter papers would be cited more than longer ones," he says. "I assumed that people don't have the time to read long papers."

Papers of about 4 pages — the length of Letters in *Astrophysical Journal* and *Astronomy and Astrophysics*, which report brief summaries of work that is usually published in more detail later — fare better than papers 5–10 pages long. But brevity offers no such benefit for papers in the other two journals considered, *Astronomical Journal* and *Monthly Notices of the Royal Astronomical Society*, which do not have Letters.

Stanek says he can't explain the length effect.

Dietrich thinks that longer papers are more useful because on average they contain more information. But he suspects that such papers are often not read in their entirety. A study of the propagation of citation errors has revealed that citations are often simply copied without being read².

Stanek does not intend to submit his study to a journal (although he encourages readers to "feel free to cite it as often as possible"). But it follows other recent investigations of citation statistics in astronomy^{3,4}, which looked at the effects of placing papers on the preprint server <http://arxiv.org>. On average this doubles the number of citations received³, and Dietrich found that the timing of posting, and thus the paper's position in the mailing list, also had an important effect⁴.

This apparent potential to rig citations

"There is definitely too much obsession with citations and other indices."

Gene chips unmask cryptic diseases

People with diverse symptoms including mental retardation, small head size, heart problems and cataracts have genomic rearrangements on the same region of chromosome 1, researchers reported last week.

A team led by genome scientist Evan Eichler at the University of Washington in Seattle linked deletions and duplications in a region of chromosome 1 that is 1.35 million DNA bases long to the abnormalities and cases of autism and learning difficulties¹. Two studies published in July linked the same genomic region to schizophrenia^{2,3}.

The findings are just the latest in a spate of studies using microarrays, which measure genetic make-up and activity, to identify small DNA defects in patients with complex disorders. Doctors are already using the technique to diagnose genetic causes of unexplained developmental conditions after clinical tests have proven inconclusive, making microarray tests an early clinical success story in the genomic revolution.

"The technology is moving rapidly from research to clinical labs, and what's research information today is becoming part of our clinical practice tomorrow," says David Ledbetter, an expert in chromosomal abnormalities at Emory University in Atlanta, Georgia.

Microarrays are glass slides embedded with fragments of DNA selected from entire genomes. Scientists can detect abnormalities in a patient's DNA — such as losses or gains of hundreds or a thousand bases — by comparing it with the normal material on the microarray. These deletions or duplications, also called copy-number changes, can change the dosage of a gene. They were invisible to older techniques

such as karyotype analysis and fluorescence *in situ* hybridization, both of which involve examining whole chromosomes.

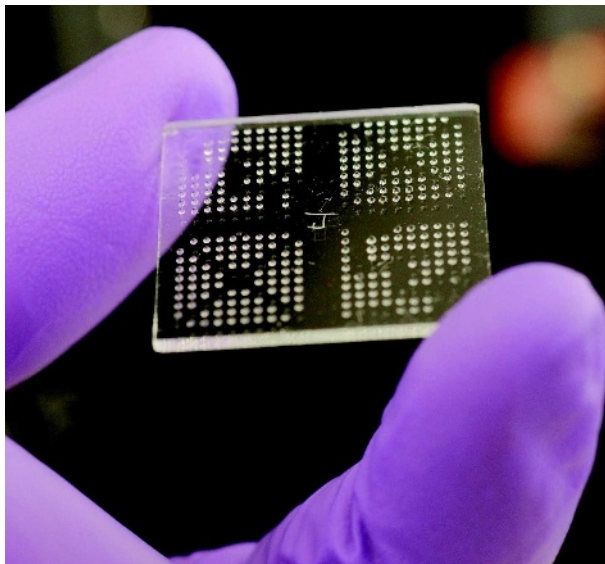
Over the past two years, microarray-based DNA comparisons have become routine in clinics around the world. Microarray supplier Agilent, based in Santa Clara, California, estimates that the worldwide market in this form of microarray analysis approaches US\$200 million and is

growing rapidly.

But a new appreciation of the amount of genetic variation found in healthy people means it can be difficult to tell whether a particular change is causing a patient's symptoms.

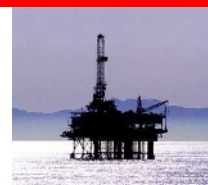
In Eichler's study, for example, eight patients had spontaneous deletions in the chromosome 1 region, but nine patients inherited their abnormalities from a parent, some of whom showed no symptoms. It's not clear how rearrangements in the region, which contains nine genes, cause the many symptoms seen in the study.

Such findings complicate the use of microarray tests in prenatal diagnostics, an area that is proceeding cautiously. On 1 October, doctors will begin enrolling 4,000 patients in a study funded by the US National Institute of Child Health and Human Development in Bethesda, Maryland, to



Microarrays are an early genomic success story — the technology is already being used in the clinic.

PHOTO: JEFFREY M. HARRIS/CHAGNON



OFFSHORE DRILLING: BLACK GOLD?

The United States debates its exploration of the sea for oil and gas.

www.nature.com/news

simply by adjusting a paper's size, place and timing of exposure sounds worrying at a time when citation statistics are increasingly being used, both formally and informally, as indicators of performance. Stanek points out that the '*h* index'⁵, a measure of the cumulative impact of a researcher's published output, has become a fetish among some scientists. "You will become obsessed with it," he advises young scientists in his paper, and will start "telling other astronomers that your *h* is bigger than their *h*".

"There is definitely too much obsession with citations and other indices," Stanek says. This is partly because they are easy to calculate regardless of whether they actually mean very much. But he confesses to using them himself.

Stanek treats his results semi-humorously, using them with earlier studies to draw up a set of guidelines for how graduate students

might manage their publications and citations to greatest advantage as their careers progress. "Make sure you submit your papers to [Arxiv] just after 4 p.m. US Eastern time on Wednesday," he suggests, for example.

"Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely."

But his study highlights some important questions. One is whether — in the face of new dissemination channels such as preprint servers and an increased sensitivity to citation indices — it is realistic to regard citations as an accurate measure of achievement.

Another question is how long a paper ought to be. If length really does matter, will that encourage researchers simply to inflate their results unnecessarily? Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely. "Most astronomical publications in the nineteenth century were very short observation reports, a few paragraphs

at most," says Dietrich. But Stanek has no problem with length, saying that he prefers papers to "be as self-contained as possible — and I have seen a lot of short papers that are not".

Yet Dietrich thinks the risk of encouraging people to inflate their papers with waffle is small. "Writing a bloated paper is considerably more work, and incurs the risk of diluting an interesting work to the point that readers don't find the interesting aspects. Also, referees and editors are usually very critical of bloat. The best tactic is still to write a good paper, and a bloated paper is not a good paper."

Philip Ball

1. Stanek, K. Z. Preprint at <http://arxiv.org/abs/0809.0692> (2008).
2. Simkin, M. V. & Roychowdhury, V. P. *Complex Syst.* **14**, 269 (2003).
3. Schwarz, G. J. & Kennicutt, R. C. Jr Preprint at <http://arxiv.org/abs/astro-ph/0411275> (2004).
4. Dietrich, J. P. Preprint at <http://arxiv.org/abs/0712.1037> (2007).
5. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* **102**, 16569–16572 (2005).

compare microarray-based tests with traditional prenatal diagnosis techniques.

Scientists worldwide are trying to decrease uncertainty by pooling their samples in databases of copy-number aberrations, including Canada's Database of Genomic Variants, DECIPHER in the United Kingdom, and the European Union-funded ECARUCA. The databases help scientists link rare DNA changes to conditions ranging from autism and schizophrenia to kidney disease.

Eventually, scientists and doctors hope to understand why changes in gene copy number cause disease. But Eichler warns that this will require studies with perhaps tens of thousands of patients, as well as consultations with patients and their families for follow-up analyses, something that isn't part of most large genome-wide association studies today.

It will also require US geneticists to share samples as freely as their European colleagues, says Eichler, who is grateful to his European collaborators, but laments his US colleagues' reluctance to pool their resources on a large scale. Ledbetter is trying to coax US geneticists to share with the help of a grant from the American College of Medical Genetics Foundation in Rockville, Maryland. Eichler hopes the plan will succeed: "We are going to need a sea change."

Erika Check Hayden

1. Mefford, H. C. *et al.* *New Engl. J. Med.* Advance online publication doi:10.1056/NEJMoa0805384 (2008).
2. Stefansson, H. *et al.* *Nature* **455**, 232–236 (2008).
3. International Schizophrenia Consortium *Nature* **455**, 237–241 (2008).

SNAPSHOT Long-lost ancestor

This pale-yellow, eyeless creature is so bizarre that naturalist E. O. Wilson named it "the ant from Mars". *Martialis heureka*, a native of the Brazilian Amazon, is the founding member of a new subfamily of ants, which split off from the ant family tree early in its evolution.

"It could represent a relict species that retained some ancestral morphological characteristics," says discoverer Christian Rabeling, a graduate

student in integrative biology at the University of Texas in Austin.

Ants evolved from wasps, so it was long assumed that any living ancestral species would be wasp-like and similar to a Cretaceous ant fossil discovered in the 1960s by Wilson and his colleagues. But *Martialis* stunned entomologists by looking completely different — genetic analysis confirms that it doesn't fit into the known taxonomy of ants (C. Rabeling *et al.* *Proc. Natl*

Acad. Sci. USA doi:10.1073/pnas.0806187105; 2008). It has long, delicate mouthparts, for munching soft invertebrates perhaps. And, compared with its sturdy front legs, the rear two sets are thin and spindly (the three other legs in the specimen shown were lopped off for DNA analysis). "It doesn't even look like it could walk at all," says Brian Fisher, an ant expert and curator of entomology at the California Academy of Sciences in San Francisco.

Amber Dance



C. RABELING/UNIV. TEXAS, AUSTIN

A longer paper gathers more citations

Researchers could garner more citations simply by making their papers longer, a study seems to imply.

In an analysis of 30,027 peer-reviewed papers published between 2000 and 2004 in top astronomy journals, astronomer Krzysztof Stanek of Ohio State University in Columbus found that the median number of citations increases with the length of the paper — from just 6 for papers of 2–3 pages to about 50 for 50-page papers¹.

There is, however, a limit to the benefits of size: citations start to tail off when papers reach lengths of 80 pages or so, perhaps because fewer people have the stamina to read them.

It is unexpected, says astronomer Jörg Dietrich of the European Southern Observatory

headquarters in Germany, who recently conducted a similar analysis and found the same results but didn't publish them. "I expected that shorter papers would be cited more than longer ones," he says. "I assumed that people don't have the time to read long papers."

Papers of about 4 pages — the length of *Letters in Astrophysical Journal* and *Astronomy and Astrophysics*, which report brief summaries of work that is usually published in more detail later — fare better than papers 5–10 pages long. But brevity offers no such benefit for papers in the other two journals considered, *Astronomical Journal* and *Monthly Notices of the Royal Astronomical Society*, which do not have Letters.

Stanek says he can't explain the length effect.

Dietrich thinks that longer papers are more useful because on average they contain more information. But he suspects that such papers are often not read in their entirety. A study of the propagation of citation errors has revealed that citations are often simply copied without being read².

Stanek does not intend to submit his study to a journal (although he encourages readers to "feel free to cite it as often as possible"). But it follows other recent investigations of citation statistics in astronomy^{3,4}, which looked at the effects of placing papers on the preprint server <http://arxiv.org>. On average this doubles the number of citations received³, and Dietrich found that the timing of posting, and thus the paper's position in the mailing list, also had an important effect⁴.

This apparent potential to rig citations

"There is definitely too much obsession with citations and other indices."

Gene chips unmask cryptic diseases

People with diverse symptoms including mental retardation, small head size, heart problems and cataracts have genomic rearrangements on the same region of chromosome 1, researchers reported last week.

A team led by genome scientist Evan Eichler at the University of Washington in Seattle linked deletions and duplications in a region of chromosome 1 that is 1.35 million DNA bases long to the abnormalities and cases of autism and learning difficulties¹. Two studies published in July linked the same genomic region to schizophrenia^{2,3}.

The findings are just the latest in a spate of studies using microarrays, which measure genetic make-up and activity, to identify small DNA defects in patients with complex disorders. Doctors are already using the technique to diagnose genetic causes of unexplained developmental conditions after clinical tests have proven inconclusive, making microarray tests an early clinical success story in the genomic revolution.

"The technology is moving rapidly from research to clinical labs, and what's research information today is becoming part of our clinical practice tomorrow," says David Ledbetter, an expert in chromosomal abnormalities at Emory University in Atlanta, Georgia.

Microarrays are glass slides embedded with fragments of DNA selected from entire genomes. Scientists can detect abnormalities in a patient's DNA — such as losses or gains of hundreds or a thousand bases — by comparing it with the normal material on the microarray. These deletions or duplications, also called copy-number changes, can change the dosage of a gene. They were invisible to older techniques

such as karyotype analysis and fluorescence *in situ* hybridization, both of which involve examining whole chromosomes.

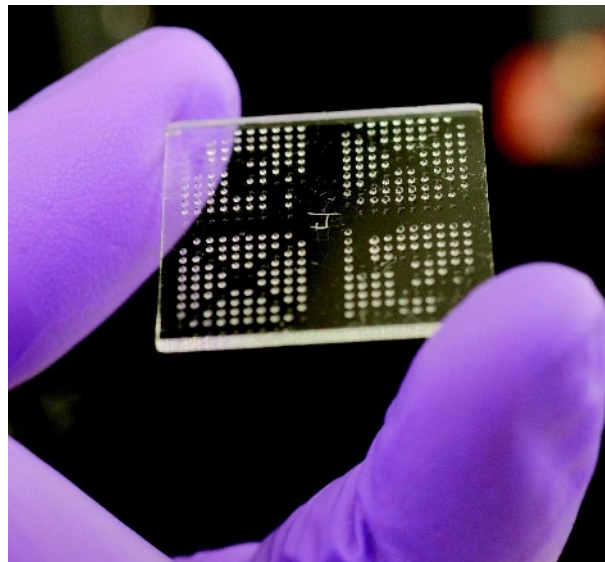
Over the past two years, microarray-based DNA comparisons have become routine in clinics around the world. Microarray supplier Agilent, based in Santa Clara, California, estimates that the worldwide market in this form of microarray analysis approaches US\$200 million and is

growing rapidly.

But a new appreciation of the amount of genetic variation found in healthy people means it can be difficult to tell whether a particular change is causing a patient's symptoms.

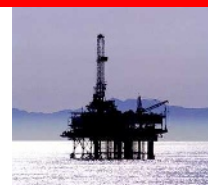
In Eichler's study, for example, eight patients had spontaneous deletions in the chromosome 1 region, but nine patients inherited their abnormalities from a parent, some of whom showed no symptoms. It's not clear how rearrangements in the region, which contains nine genes, cause the many symptoms seen in the study.

Such findings complicate the use of microarray tests in prenatal diagnostics, an area that is proceeding cautiously. On 1 October, doctors will begin enrolling 4,000 patients in a study funded by the US National Institute of Child Health and Human Development in Bethesda, Maryland, to



Microarrays are an early genomic success story — the technology is already being used in the clinic.

PHOTO: JEFFREY M. HARRIS/CHAGNON



OFFSHORE DRILLING: BLACK GOLD?

The United States debates its exploration of the sea for oil and gas.

www.nature.com/news

simply by adjusting a paper's size, place and timing of exposure sounds worrying at a time when citation statistics are increasingly being used, both formally and informally, as indicators of performance. Stanek points out that the '*h* index'⁵, a measure of the cumulative impact of a researcher's published output, has become a fetish among some scientists. "You will become obsessed with it," he advises young scientists in his paper, and will start "telling other astronomers that your *h* is bigger than their *h*".

"There is definitely too much obsession with citations and other indices," Stanek says. This is partly because they are easy to calculate regardless of whether they actually mean very much. But he confesses to using them himself.

Stanek treats his results semi-humorously, using them with earlier studies to draw up a set of guidelines for how graduate students

might manage their publications and citations to greatest advantage as their careers progress. "Make sure you submit your papers to [Arxiv] just after 4 p.m. US Eastern time on Wednesday," he suggests, for example.

"Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely."

But his study highlights some important questions. One is whether — in the face of new dissemination channels such as preprint servers and an increased sensitivity to citation indices — it is realistic to regard citations as an accurate measure of achievement.

Another question is how long a paper ought to be. If length really does matter, will that encourage researchers simply to inflate their results unnecessarily? Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely. "Most astronomical publications in the nineteenth century were very short observation reports, a few paragraphs

at most," says Dietrich. But Stanek has no problem with length, saying that he prefers papers to "be as self-contained as possible — and I have seen a lot of short papers that are not".

Yet Dietrich thinks the risk of encouraging people to inflate their papers with waffle is small. "Writing a bloated paper is considerably more work, and incurs the risk of diluting an interesting work to the point that readers don't find the interesting aspects. Also, referees and editors are usually very critical of bloat. The best tactic is still to write a good paper, and a bloated paper is not a good paper."

Philip Ball

1. Stanek, K. Z. Preprint at <http://arxiv.org/abs/0809.0692> (2008).
2. Simkin, M. V. & Roychowdhury, V. P. *Complex Syst.* **14**, 269 (2003).
3. Schwarz, G. J. & Kennicutt, R. C. Jr Preprint at <http://arxiv.org/abs/astro-ph/0411275> (2004).
4. Dietrich, J. P. Preprint at <http://arxiv.org/abs/0712.1037> (2007).
5. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* **102**, 16569–16572 (2005).

compare microarray-based tests with traditional prenatal diagnosis techniques.

Scientists worldwide are trying to decrease uncertainty by pooling their samples in databases of copy-number aberrations, including Canada's Database of Genomic Variants, DECIPHER in the United Kingdom, and the European Union-funded ECARUCA. The databases help scientists link rare DNA changes to conditions ranging from autism and schizophrenia to kidney disease.

Eventually, scientists and doctors hope to understand why changes in gene copy number cause disease. But Eichler warns that this will require studies with perhaps tens of thousands of patients, as well as consultations with patients and their families for follow-up analyses, something that isn't part of most large genome-wide association studies today.

It will also require US geneticists to share samples as freely as their European colleagues, says Eichler, who is grateful to his European collaborators, but laments his US colleagues' reluctance to pool their resources on a large scale. Ledbetter is trying to coax US geneticists to share with the help of a grant from the American College of Medical Genetics Foundation in Rockville, Maryland. Eichler hopes the plan will succeed: "We are going to need a sea change."

Erika Check Hayden

1. Mefford, H. C. *et al.* *New Engl. J. Med.* Advance online publication doi:10.1056/NEJMoa0805384 (2008).
2. Stefansson, H. *et al.* *Nature* **455**, 232–236 (2008).
3. International Schizophrenia Consortium *Nature* **455**, 237–241 (2008).

SNAPSHOT

Long-lost ancestor

This pale-yellow, eyeless creature is so bizarre that naturalist E. O. Wilson named it "the ant from Mars". *Martialis heureka*, a native of the Brazilian Amazon, is the founding member of a new subfamily of ants, which split off from the ant family tree early in its evolution.

"It could represent a relict species that retained some ancestral morphological characteristics," says discoverer Christian Rabeling, a graduate

student in integrative biology at the University of Texas in Austin.

Ants evolved from wasps, so it was long assumed that any living ancestral species would be wasp-like and similar to a Cretaceous ant fossil discovered in the 1960s by Wilson and his colleagues. But *Martialis* stunned entomologists by looking completely different — genetic analysis confirms that it doesn't fit into the known taxonomy of ants (C. Rabeling *et al.* *Proc. Natl*

Acad. Sci. USA doi:10.1073/pnas.0806187105; 2008). It has long, delicate mouthparts, for munching soft invertebrates perhaps. And, compared with its sturdy front legs, the rear two sets are thin and spindly (the three other legs in the specimen shown were lopped off for DNA analysis). "It doesn't even look like it could walk at all," says Brian Fisher, an ant expert and curator of entomology at the California Academy of Sciences in San Francisco.

Amber Dance



C. RABELING/UNIV. TEXAS, AUSTIN



OFFSHORE DRILLING: BLACK GOLD?

The United States debates its exploration of the sea for oil and gas.

www.nature.com/news

simply by adjusting a paper's size, place and timing of exposure sounds worrying at a time when citation statistics are increasingly being used, both formally and informally, as indicators of performance. Stanek points out that the '*h* index'⁵, a measure of the cumulative impact of a researcher's published output, has become a fetish among some scientists. "You will become obsessed with it," he advises young scientists in his paper, and will start "telling other astronomers that your *h* is bigger than their *h*".

"There is definitely too much obsession with citations and other indices," Stanek says. This is partly because they are easy to calculate regardless of whether they actually mean very much. But he confesses to using them himself.

Stanek treats his results semi-humorously, using them with earlier studies to draw up a set of guidelines for how graduate students

might manage their publications and citations to greatest advantage as their careers progress. "Make sure you submit your papers to [Arxiv] just after 4 p.m. US Eastern time on Wednesday," he suggests, for example.

"Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely."

But his study highlights some important questions. One is whether — in the face of new dissemination channels such as preprint servers and an increased sensitivity to citation indices — it is realistic to regard citations as an accurate measure of achievement.

Another question is how long a paper ought to be. If length really does matter, will that encourage researchers simply to inflate their results unnecessarily? Many of the greatest past discoveries, such as the structure of DNA, were reported extremely concisely. "Most astronomical publications in the nineteenth century were very short observation reports, a few paragraphs

at most," says Dietrich. But Stanek has no problem with length, saying that he prefers papers to "be as self-contained as possible — and I have seen a lot of short papers that are not".

Yet Dietrich thinks the risk of encouraging people to inflate their papers with waffle is small. "Writing a bloated paper is considerably more work, and incurs the risk of diluting an interesting work to the point that readers don't find the interesting aspects. Also, referees and editors are usually very critical of bloat. The best tactic is still to write a good paper, and a bloated paper is not a good paper."

Philip Ball

1. Stanek, K. Z. Preprint at <http://arxiv.org/abs/0809.0692> (2008).
2. Simkin, M. V. & Roychowdhury, V. P. *Complex Syst.* **14**, 269 (2003).
3. Schwarz, G. J. & Kennicutt, R. C. Jr Preprint at <http://arxiv.org/abs/astro-ph/0411275> (2004).
4. Dietrich, J. P. Preprint at <http://arxiv.org/abs/0712.1037> (2007).
5. Hirsch, J. E. *Proc. Natl Acad. Sci. USA* **102**, 16569–16572 (2005).

compare microarray-based tests with traditional prenatal diagnosis techniques.

Scientists worldwide are trying to decrease uncertainty by pooling their samples in databases of copy-number aberrations, including Canada's Database of Genomic Variants, DECIPHER in the United Kingdom, and the European Union-funded ECARUCA. The databases help scientists link rare DNA changes to conditions ranging from autism and schizophrenia to kidney disease.

Eventually, scientists and doctors hope to understand why changes in gene copy number cause disease. But Eichler warns that this will require studies with perhaps tens of thousands of patients, as well as consultations with patients and their families for follow-up analyses, something that isn't part of most large genome-wide association studies today.

It will also require US geneticists to share samples as freely as their European colleagues, says Eichler, who is grateful to his European collaborators, but laments his US colleagues' reluctance to pool their resources on a large scale. Ledbetter is trying to coax US geneticists to share with the help of a grant from the American College of Medical Genetics Foundation in Rockville, Maryland. Eichler hopes the plan will succeed: "We are going to need a sea change."

Erika Check Hayden

1. Mefford, H. C. *et al.* *New Engl. J. Med.* Advance online publication doi:10.1056/NEJMoa0805384 (2008).
2. Stefansson, H. *et al.* *Nature* **455**, 232–236 (2008).
3. International Schizophrenia Consortium *Nature* **455**, 237–241 (2008).

SNAPSHOT Long-lost ancestor

This pale-yellow, eyeless creature is so bizarre that naturalist E. O. Wilson named it "the ant from Mars". *Martialis heureka*, a native of the Brazilian Amazon, is the founding member of a new subfamily of ants, which split off from the ant family tree early in its evolution.

"It could represent a relict species that retained some ancestral morphological characteristics," says discoverer Christian Rabeling, a graduate

student in integrative biology at the University of Texas in Austin.

Ants evolved from wasps, so it was long assumed that any living ancestral species would be wasp-like and similar to a Cretaceous ant fossil discovered in the 1960s by Wilson and his colleagues. But *Martialis* stunned entomologists by looking completely different — genetic analysis confirms that it doesn't fit into the known taxonomy of ants (C. Rabeling *et al.* *Proc. Natl*

Acad. Sci. USA doi:10.1073/pnas.0806187105; 2008). It has long, delicate mouthparts, for munching soft invertebrates perhaps. And, compared with its sturdy front legs, the rear two sets are thin and spindly (the three other legs in the specimen shown were lopped off for DNA analysis). "It doesn't even look like it could walk at all," says Brian Fisher, an ant expert and curator of entomology at the California Academy of Sciences in San Francisco.

Amber Dance



C. RABELING/UNIV. TEXAS, AUSTIN

Virgo's vacuum failure stalls gravity-wave hunt

The Virgo gravity-wave interferometer, an €80-million (US\$114-million) experiment located near Pisa, Italy, has been incapacitated by a vacuum failure for most of the summer and is expected to stay out of commission for another few months.

During a test on 9 May a glass viewport in the L-shaped detector shattered, sending shards of glass into one of the laser-reflecting mirrors and its ancillary instrumentation. The mirror was damaged beyond repair and is being replaced.

The team has decided to replace all 100 viewports to avoid another failure. Francesco Fidecaro, Virgo's spokesman, who is a physicist at the University of Pisa, says the original viewports might not have been "appropriate" for the operating conditions. Physicists believe that it was not the vacuum itself but the vibrations created by the pumps that caused the port to fail. They are confident that the new viewports will not fail in the same way.

For a longer version of this story, see <http://tinyurl.com/5dfzug>

Fossil hunter accused of stealing dinosaur

A high-profile fossil collector from Montana will go to court on 23 September for allegedly stealing a US\$400,000 raptor specimen from private land.

Nate Murphy, a former curator of palaeontology at the Phillips County Museum in Malta, was charged last week with the theft of a fox-sized raptor, nicknamed 'Sid Vicious', from the 77-million-year-old Judith River Formation. State and federal agents began a probe in spring 2007 after authorities heard of a possible misappropriation from participants in Murphy's collecting business, the Judith River Dinosaur Institute, now in Billings, Montana.



Dinosaur collector Nate Murphy.

Bisphenol A exposure may pose disease risk

High levels of bisphenol A, a chemical found in some food and drinks containers, may be associated with an increased risk of cardiovascular disease and diabetes.

An epidemiological analysis of 1,455 adults in the United States, published this week (I. A. Lang *et al.* *J. Am. Med. Assoc.* **300**, 1303–1310; 2008), is the first to find a link between bisphenol A and disease in humans. The results are likely to fuel further debate over whether the chemical should be banned, but the authors caution that the study must first be replicated.

Earlier this year, the US National Toxicology Program expressed "some concern" about the possible impact of the compound on fetuses, infants and children.

In a draft report released in August, the US Food and Drug Administration stated that "food contact" products containing bisphenol A are safe. The agency has said it will continue to review the compound as more data become available.



On 14 September, Murphy and another dinosaur he helped to collect — 'Leonardo', a mummified hadrosaur with skin and organs — were featured in a Discovery Channel film on US television.

Blandford to lead next decadal astronomy survey

The US National Academies have appointed Roger Blandford to chair the next decadal survey in astronomy, starting a two-year process that culminates in a priority list of astronomy projects for the next ten years (see *Nature* **443**, 386–389; 2006).

Blandford, director of the Kavli Institute for Particle Astrophysics and Cosmology at Stanford University in California, will assemble a committee of about 20 people. They will gather input from hundreds of panel members into a final report and wish-list, expected by mid-2010. NASA, the National Science Foundation and Congress often cite the report in setting science budgets.

But the last survey, in 2001, was over-ambitious in recommending seven major projects. Cost overruns mean that none of them will be finished this decade. That survey's highest priority, the \$4.5-billion James Webb Space Telescope, is due for launch in 2013.

For a longer version of this story, see <http://tinyurl.com/6hebbqn>

US health agency gets creative with grants

The National Institutes of Health (NIH) plans to spend more than \$250 million over the next five years to inject a mega-dose of creativity into its basic grants.

Beginning with funding of \$25 million

in 2009, conventional R01 grants will be supplemented by 'transformational' awards, or T-R01s. The agency will test new peer-review approaches to pick out "bold and groundbreaking" proposals that "reflect an exceptional level of creativity" and may "promote radical changes" in a field, the NIH said last week.

Investigators working in any NIH-funded discipline can apply for the grants. The agency says that areas of creative need include the science of behaviour change; protein capture; functional variation in mitochondria; complex three-dimensional tissue models; and pharmacogenomics.

Court fines entomologists for illegal collecting in India

The case of two Czech entomologists arrested in India took a further turn last week when a court in Darjeeling sentenced Emil Kučera to three years in prison, plus a 60,000-rupee (US\$1,300) fine, under the country's Wildlife Act and Biological Diversity Act. His colleague, Petr Švácha of the Institute of Entomology of the Academy of Sciences of the Czech Republic in České Budějovice, was fined 20,000 rupees.

The pair have been held in India since 22 June for collecting butterflies from a national park without a permit. Kučera has been given bail and plans to appeal; Švácha is staying in India to help Kučera's legal fight.

"This case points to the need for revision of India's conservation legislation," says Priyadarsanan Dharma Rajan, a taxonomist with the Ashoka Trust for Research in Ecology and the Environment in Bangalore. "There are no cases of any insect gone extinct due to collection. Threats to insects come from pesticides poured into our fields and habitat loss due to land-use changes."

The heart of the wood

Białowieża is one of the best-preserved woodlands in Europe. But is it a good reference point for what Europe looked like 5,000 years ago? **Emma Marris** goes deep into the forest to find out.

In dappled light, the trunks of lindens, Norway spruces, and oaks prise the undergrowth and the canopy apart. Quite a few are dead, and draped with beards of moss. Birds sing in the cool, green-smelling air. It is very still, and at first feels almost timeless. But time is here — time past and time present. Hornbeam limbs, once weighed down with snow, are still bowed at oblique angles in the summer. Red deer bones on the ground mark an old kill; saplings clamber like children on the body of their recently fallen parent. As the wind ruffles the canopy, patches of sun traipse to and fro on the forest floor. Fat, desultory mosquitoes drift through the air.

Behind a shrub stands a European bison, or wisent, in the prime of his life, with curled black horns, a high woolly shoulder and a ridged backbone. He is browsing noiselessly. In a few minutes he will glide away as noiselessly as if he were slipping back 1,000 years to when his kind roamed forests like these over most of Europe.

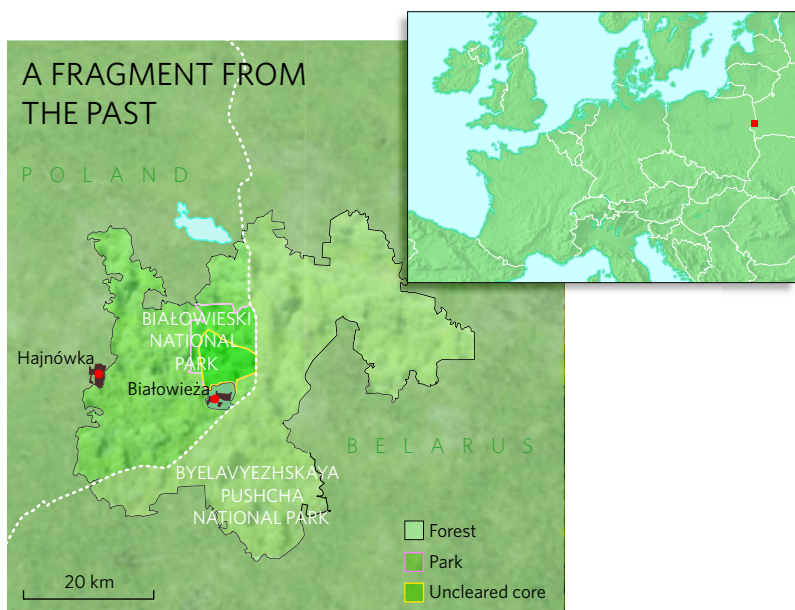
The Białowieża forest in eastern Poland is a place apart, and of the past. Grey wolves hunt red deer; woodpeckers — all ten of Europe's species live here — hunt for beetles in 500-year-old oaks; stripy young boar splash through alder bogs blooming with yellow irises. The forest spreads over some 146,000 hectares in Poland and Belarus, various parts of it are managed in various ways, and have been cleared at various times. But it is said that the 4,438 hectares at its Polish core have never been cleared. Many of the forest's ecological processes have run uninterrupted since the last ice age, and the trees have never been selected for wood that can be sawn into long, knotless boards. It is thus a reference point, showing how humans have changed forest composition and processes elsewhere. It is a model for silviculturalists interested in conservation — or in commerce, as tree plantations that mimic nature may produce more wood.

The University of Warsaw's Geobotanical Station and the Polish Academy of Sciences' Mammal Research Institute, both in Białowieża village, study the surrounding forest to get an idea of how things once were. The Mammal Research Institute is most famous for the discovery by one of its staff of the Dehnel effect, in which shrews actually shrink their skeletons to cope with reduced food supply in winter. The forest is indubitably a good place to study species in conditions most like those to which they were adapted. Ornithologists, in particular, love the forest and have noted that bird life there is diverse, but not dense; the reverse of what's seen in agricultural landscapes. But to what extent is Białowieża, as often claimed, Europe's last primeval broadleaf forest?

In his 1996 book *Natural Woodland* (Cambridge Univ. Press), the British naturalist George Peterken defines three types of naturalness. He contrasts "original-naturalness" — "the state which existed before people became a significant ecological factor" — with "present-naturalness", "the state which would prevail now if people had not become



N. BENNIE/NATUREPL.COM



a significant factor". To see the difference, remember that forests are constantly shaped by fire and storm. In the clearing, seedlings compete for light. Trees fill in the gaps. Soil is depleted and fertilized; climate alters. A forest that has travelled thousands of years down a human-free road will change on the way, as original-naturalness becomes present-naturalness. And both are distinct from "future-naturalness", "the state which would eventually develop if people's influence were completely and permanently removed".

Few environmentalists have adopted Peterken's distinctions. They prefer terms such as 'virgin' and 'old growth' forests, seeing the human factor as more crucial than any other. The problem with this, Peterken notes, is that if human use is taken in its broadest sense — not just logging — there is no more virgin forest in North America than there is in Europe. Native Americans managed their woodlands for game and maintained open spaces using fire.

Royal playground

Białowieża cannot aspire to present-naturalness. It owes its existence to the Lithuanian dukes, Polish kings, Russian tsars and German Nazis who kept it to themselves for hunting, first to feed their armies, later for sport. The forest ungulates were fed winter hay. Competing predators were killed; there are no more bears in Białowieża. The European mink that once flourished are also gone, their place now taken — apparently without much ecological disruption — by Asian raccoon dogs and American mink escaped from fur farms. Before they went extinct in the early seventeenth century, aurochs, the ancestor to domestic cattle, probably shared the forest with wisent and red and roe deer. An unknown poacher shot the forest's last wild bison in April 1919; today's are descended from 13 zoo bison, the descendants of which were reintroduced into the wild in the 1950s.

This history of intervention makes the forest's ecology hard to read. Take the recent reduction in Norway spruce. Their share of the woodland, says Bogdan Jaroszewicz, head of the Geobotanical Station, has halved since 1950. Some believe the spruce trees are disappearing because of climate

change; the park is at the southern limit of the spruce's range. Others blame the spruce bark beetle, although its effects may be synergistic with, not alternative to, those of climate change. But Jaroszewicz points out that a century ago Russian royalty were breeding so much game in the woods that very few saplings survived to maturity. "There was almost no regeneration," he says. This intense browsing might have had a greater impact on seedlings more succulent than those of spruce. No one can guess whether the spruce would be declining without previous human intervention.

What then of the possibility of original-naturalness — the primeval state preserved? This is certainly the pitch of Białowieża's growing tourist trade. Definitions of 'natural' or 'pristine' or 'primeval' forest typically include trees of many different ages coexisting and the presence of marker species dependent on old or dead wood — often mosses, lichens and fungi. Białowieża has all these characteristics. Trees age, die and rot where they fall, fuelling the life cycles of innumerable beetles, fungi, woodpeckers and, eventually, the next generation of trees. In Białowieża's core up to one quarter, by volume, of the wood above ground in each hectare is dead, compared with 2% in commercial stands.

Strange forces

People have interacted with Białowieża for millennia, but not necessarily as agents of change — their needs remained constant. Hunting has affected the numbers of large animals, but pollen records suggest that in terms of the species that make up the forest's main constituents, Białowieża has stayed fairly constant for millennia. Tomasz Samojlik, who grew up near the forest, has returned to study its environmental history at the Polish Academy of Sciences' Mammal Research Institute in Białowieża village. He has investigated royal hunts, beekeeping in trees, ore-smelting operations deep in the forest, and more. "After five or six years I have come to the conclusion that humans have always been connected to the forest," he says. For humans, he adds, the forest could be a matter of life and death: "They were part of it and could be killed by it easily." But for the forest, humans were one factor among many — they shaped it in some ways, left it alone in others, and were no threat to it at all.

That was then. "Now we are some strange force from an outside world," says Samojlik. All concerned want to limit that force's impact. The humans who now manage Białowieża struggle to do so in a way that leaves no fingerprints, aspiring, in Peterken's terms, to something close to the future-natural. Tourists may enter only when accompanied by guides. A fence, built in the 1930s to exclude domestic animals, now stands as a barrier between the domains of humans and the woods. Passing through sunny fields to reach a shaded gate into the forest, researchers mark the separation of realms by anointing themselves with mosquito repellent.

But future-natural takes more than a fence. The Geobotanical Station's records show rum goings-on over the past 50 years that look like the work of climate: plants flowering earlier, changes in annual distribution of precipitation, changes in snowmelt timing. In a world where today's saplings will experience an atmosphere containing twice as much carbon dioxide as their parents' generation, a human-free future-naturalness is not an option.

A local approximation to future-naturalness may



Flora v. Fauna

Today there is virtually no management intervention in the Polish core of Białowieża. But the Parks Department, foresters and scientists disagree on how to manage the rest of the forest. The Parks Department intervenes more than the researchers at the Mammal Institute would recommend, feeding the bison in winter, for example, and culling them in summer.

"Most managers believe there are too many bison," says Rafał Kowalczyk of the Mammal Institute as he races a truck down a forest road in pursuit of the signal from a radio-collared bison. (Over short distances he can track his charges by smell: "Not a smell as strong as wild boar or red deer. A bit like cattle; the smell of fresh milk.") Kowalczyk wants to stop culling and reduce winter feeding. "They survived for 1,000 years

without winter feeding," he says, "and I think densities could be higher." Małgorzata Karaś, director of Białowieża National Park disagrees, saying that the culling is done "according to the habitat capacity".

Kowalczyk thinks this habitat capacity is arbitrary. "There was a meeting in the 1960s and they

came up with 250 bison," he says. "It was a number taken from heaven." But foresters worry that more bison — the target is now 300 — would damage trees and prevent regeneration in parts of the forest used for timber.

Bogumiła Jędrzejewska, who joined the mammal institute in 1985, says that communication

between various parties has improved as foresters and local people see the value of the natural forest — usually expressed in tourist revenue. "Białowieża changed its opinion," she says. "And the attitude of foresters has changed very much. It has become unpopular to focus only on timber protection." But some habits die hard. "Foresters," says Jędrzejewska, "believe that the forest wouldn't survive without their intervention, and the main intervention is through the axe."

What unites the Poles is suspicion of their neighbours. Belarusian president Alexander Lukashenko, has, says Jędrzejewska, carved a hideaway out of his country's part of the forest, complete with rollerblading path. "Formally it is protected," says Jędrzejewska. "In reality it is logged."

E.M.



AFTER THE HUNT IN THE FOREST IN 1409 V. I. NAVOZOV

sometimes be possible. Unfortunately, though, Białowieża faces a further handicap in this regard. To deal with change, an ecosystem needs space — to provide refuges for the species that will recolonize fire and wind clearings, room for ranges to change, and so on. Białowieża is quite small — only big enough, for example, for the 100-square-kilometre range of a single male lynx. "We have three or four, because ranges overlap, but hardly enough to sustain a population," says Jaroszewicz. The minimum size of a functioning forest probably varies greatly, but a calculation for Canadian boreal forest suggested about 650,000 hectares as the minimum size (S. J. Leroux *et al. Biol. Conserv.* **138**, 464–473; 2007).

Here and now

But if neither present-natural nor original-natural, and with poor odds for becoming future-natural, Białowieża is still the best we have. For researchers looking for a proxy for European forests unchanged by modern man, Białowieża is a starting point; they still have to comb out the possible effects of human activity; to sort through tangled puzzles such as the spruce decline; to account for the small size of the forest and the changes in megafauna; and to race against the clock as climate change rearranges relationships.

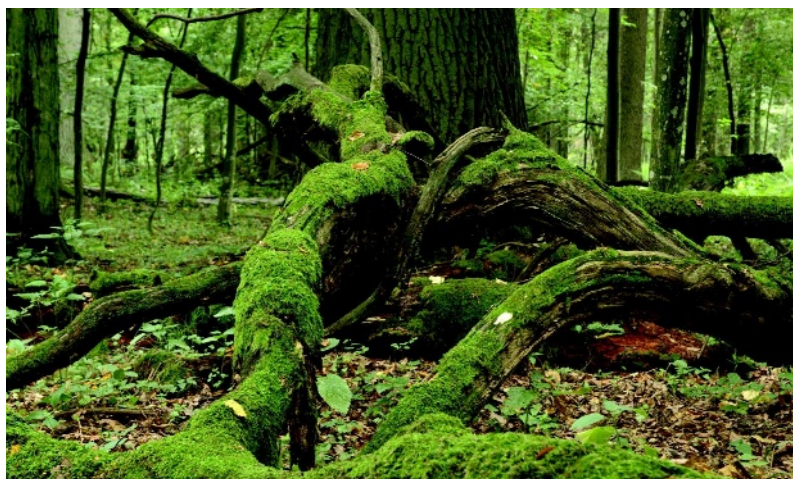
In the end, it is the biological relationships which constitute the forest that are important to the science of ecology, not what history says about 'touched' and 'untouched'. History is uncertain: easily lost, written by victors, disproportionately concerned with the urban rich. Biology is available now, changing but measurable, and Białowieża has it in abundance. "Even if we estimate that 5,000 papers have been written about it, we still only know a little about it," says Jaroszewicz. "We still don't know what it will be in 100 years." The number of species, the complex interactions researchers

track from year to year, and the individual trees, among the oldest and tallest in Europe, have a value beyond any question of whether they have been interfered with or not.

Wonder — formalized in science and experienced directly — trumps purity when it comes to feeling natural. Białowieża has long been human — but in a special and timeless way. Once, Samojlik found a sickle knife dropped by a previous visitor to the wood lying just under the tangled grass; it had been there for 1,500 years. Moss-hung and teeming with life, Białowieża retains one of the essential features of the European forest, at least as popularly imagined. It is still profoundly mysterious, even if it is no longer trackless or boundless.

Emma Marris writes for *Nature* from Columbia, Missouri.

See Editorial, page 263.



M. SZYMURA

Postmodern evolution?

This summer a group of high-profile researchers met in Altenberg, Austria, to try and plot the future course of evolutionary theory. **John Whitfield** was there.

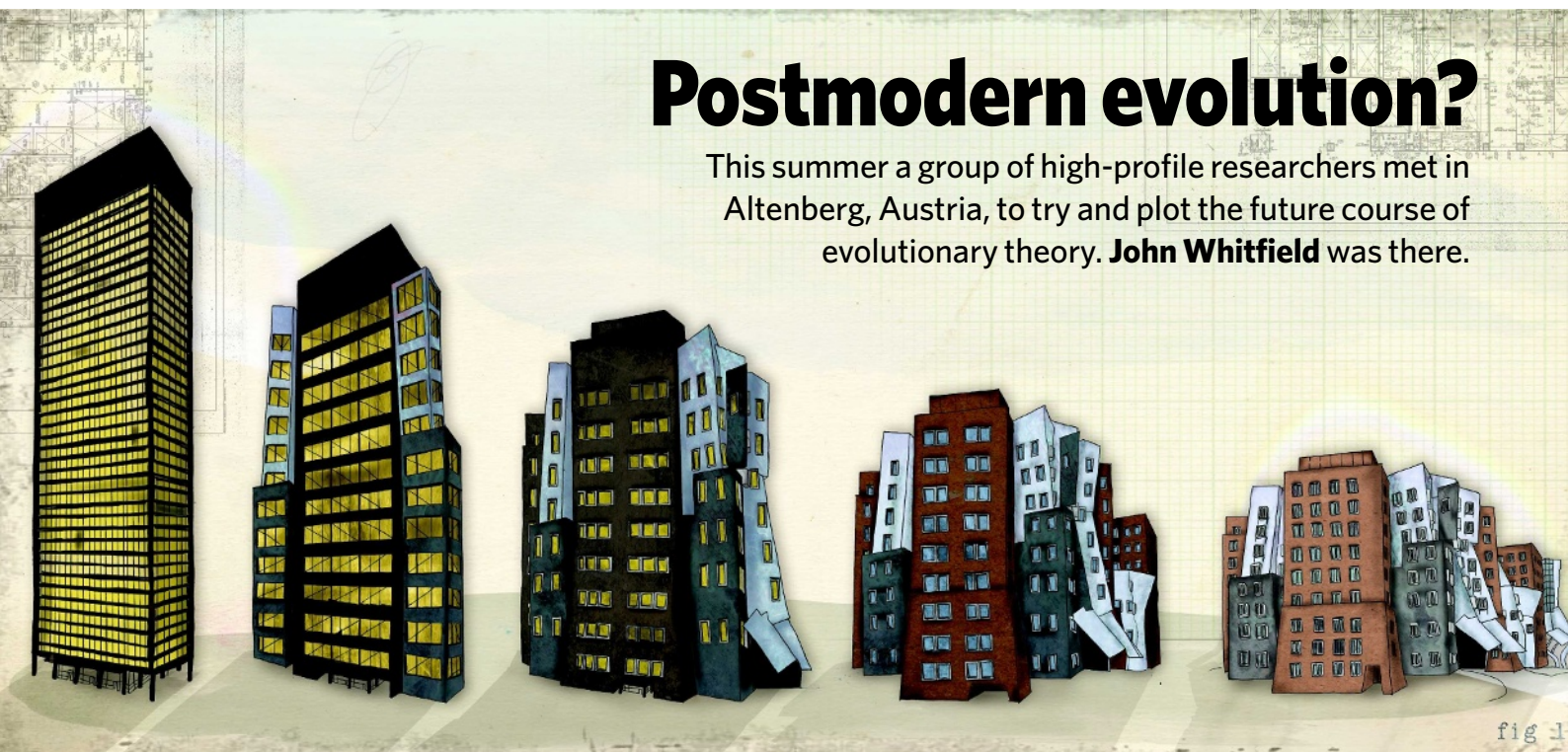


fig 1

"Oh my gosh," says Massimo Pigliucci, "maybe I shouldn't use that term." Pigliucci, responding to comments on his talk about how living things respond to their environment, and what it means for evolution, has just let slip the p-word. Later the same day, Günter Wagner, an evolutionary theorist at Yale University in New Haven, Connecticut, puts up a slide bearing the words 'Postmodern Synthesis'. Pigliucci is moved to make an editorial suggestion from the floor: "I'd really rather we didn't use that term." Wagner says the slide was intended to be tongue-in-cheek, but Pigliucci is worried about the impression the word creates: "If there's one thing we don't want, it's for people to get the idea that there's a bunch of evolutionary theories out there, and that they're all equal."

A lot of scientists loathe what they take to be postmodernism's intellectual relativism, and shy away from using the word. But doing so puts Pigliucci in something of a bind. An evolutionary ecologist at the State University of New York in Stony Brook, Pigliucci is one of the conveners of this small meeting on the future of evolutionary thought taking place at the Konrad Lorenz Institute for Evolution and Cognition Research in Altenberg, Austria. The meeting has received a fair amount of hype — in the blogosphere it was dubbed 'The Woodstock of Evolution'. Its agenda is, pretty explicitly, to go beyond the 'modern synthesis' that has held sway in evolutionary theory since

the middle of the twentieth century. And in everyday speech, it is pretty clear what comes after the modern.

What's more, some of this work sounds as though it fits the term quite nicely. Over dinner at the meeting's end, Pigliucci expresses his hope of "moving from a gene-centric view of causality in evolution to a pluralist, multi-level causality". Postmodernists in the humanities call this 'decentering', and they are all for it. Over the course of the meeting, it's fairly clear that the means to this pluralist end are being sought through mixing and matching neglected ideas and old problems from biology's past with the latest experimental and analytical techniques. Apply that sort of bricolage to architecture and you get the sort of brutalist-right-angle here, classical-column-there, swirling-titanium-ceiling-above-it-all look that is normally pigeonholed, for better or worse, as postmodern.

Evolution of ideas

Leaving aside the troublesome adjective, what is the modernism that the Altenberg meeting is meant to move beyond — or to use Pigliucci's preferred term, 'extend'? Between about 1920 and 1940, researchers such as the American Sewall Wright and the Englishmen Ronald Fisher and J. B. S. Haldane took Charles

Darwin's ideas about natural selection and Gregor Mendel's insights into how traits pass from parents to offspring — which many biologists of the time believed antithetical — and fused them into a mathematical description of the genetic makeup of populations and how it changes. That fusion was the modern synthesis. It treats an organism's form, or phenotype, as a readout of its hereditary information, or genotype. Change is explained as one version of a

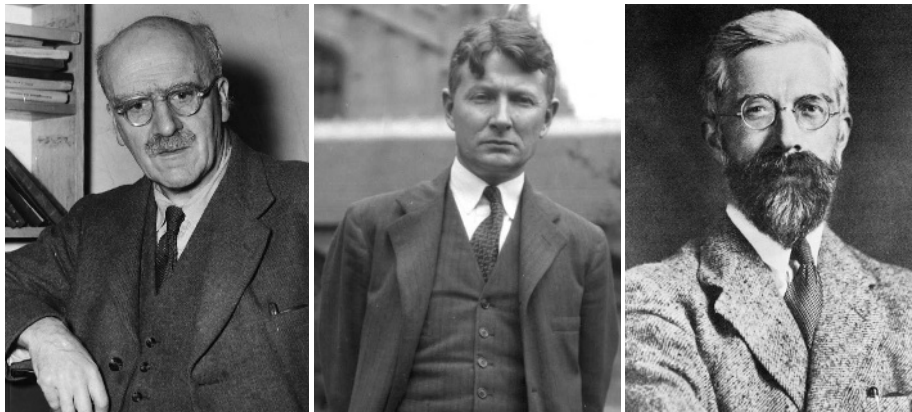
gene being replaced by another. Natural selection acts by changing the frequency of genes in the next generation according to the fitness of phenotypes in this one. In this world view, the gene is a black box, its relationship to phenotype is a one-way street, and the environment, both cellular and external, is a selective filter imposed on the readout of the genes, rather

than something that can influence an organism's form directly.

What's wrong with this picture, say the would-be extenders at Altenberg and elsewhere, is what it leaves out. Molecular biology, cell biology and genomics have provided a much richer picture of how genotypes make phenotypes. The extenders claim that enough insights have now come from this and other research for it to be time to re-examine problems that the modern synthesis doesn't address. These problems include some of the key turning

"The modern synthesis is good at modelling the survival of the fittest, but not the arrival of the fittest."
— Scott Gilbert

ILLUSTRATION BY J. TAYLOR



Modern synthesisists: (left to right) J. B. S. Haldane, Sewall Wright and Ronald Fischer.

points in evolution: the patterns and changes seen in the fossil record as new branches spring from the tree of life and new anatomies — skeletons, limbs, brains — come into being. “When the public thinks about evolution, they think about the origin of wings and the invasion of the land,” says Graham Budd, a palaeobiologist at the University of Uppsala, Sweden. “But these are things that evolutionary theory has told us little about.”

Bring on the kangaroos

The question of how form changes in individuals is the province of developmental biology, and genetic studies have now revealed a lot about how the mechanisms of development have evolved. Many see the evolutionary developmental biology — ‘evo-devo’ — that is emerging from this work as the key ingredient needed to extend or surpass the modern synthesis.

“Evolution needs a theory of body construction and change, as well as population construction and change,” says Scott Gilbert, an evo-devo researcher at Swarthmore College in Pennsylvania, who was not in Altenberg but who is writing a book on extending the evolutionary synthesis in similar directions. “The modern synthesis is remarkably good at modelling the survival of the fittest, but not good at modelling the arrival of the fittest.” To explain the production of novel features, such as limbs and feathers, Gilbert and like-minded biologists want a theory in which the environment is defined broadly enough to include the developing body, which is the primary context in which the genes are expressed. Genes shape this developing environment, but the dynamic environment also shapes the expression of the genes. And it does so directly, rather than through some later selection. “The gene will continue to be centre stage,” says Gilbert, “but it will be seen as both active and acted upon. It’s not going to be the unmoved mover.”

The importance of the environment acting

on the genome can be seen in plasticity, the ability of the same genes to give rise to radically different phenotypes in different conditions — as studied by several of the Altenberg group. Pigliucci, who works on invasive plant species, gave the example of species that lie low in a new environment for several years before becoming a problem. He puts this down to plasticity and the Baldwin effect. In 1896 James Baldwin, an American psychologist, suggested that over the generations, tricks that at first have to be learned can become hard-wired as genes fix variations caused by the environment. “It could be that the plants arrive in a new environment and hang on thanks to plasticity — it gains time for natural selection to kick in,” says Pigliucci. To begin with, the genes follow adaptation rather than leading it, as “bookkeepers of what’s happening”. Once the genes have caught up, and the immigrant can take adaptation to the environment as read, it is able to become dominant.

Plasticity also allows organisms to make the most of their mutations. “The myopic view — that we don’t need to worry about phenotypic variation, that it is abundant, always small and that it goes in all possible directions — doesn’t correspond to the conservation we’ve seen in developmental systems,” Marc Kirschner, a systems biologist at Harvard University in Cambridge, Massachusetts, told the Altenberg meeting. To grow a limb you don’t need mutations in every gene involved in limb building; life can use the facts that muscle cells naturally align with bone, nerve cells stabilize when they plug into muscles, and blood vessels grow towards areas low in oxygen to leverage

a small genetic change into an important difference. Again, the changing environment within the developing body is part of the process by which the gene is expressed: Kirschner calls it facilitated variation².

As an example, he points to the discovery that the narrow, tweezer-like beak of an insect-eating finch can become the fat, nutcracking beak of a seed-eater by increasing the activity of a single gene involved in bone formation³. “Because developmental systems are so integrated and self-regulating, you can make a large functional change without a large genetic change,” says Kirschner. Pigliucci gave a more speculative example of the possible evolutionary consequences of such changes, showing a slide juxtaposing a kangaroo and a dog that had been born without forelimbs but learnt to walk on its hind legs. “It’s hard to imagine that this kind of change doesn’t have anything to do with the evolution of bipedalism,” he told the meeting.

Self-organizing cells

Pigliucci and Kirschner think that the capacity of small genetic changes to trigger large shifts results in waves of innovation separated by seeming lulls in which evolution stabilizes and integrates the new arrangements. This matches some aspects of the fossil record, where bursts of innovation and diversification are interspersed by much longer periods of stasis — a pattern known as punctuated equilibrium, first described by the late Stephen Jay Gould and Niles Eldredge of the American Museum of Natural History in the 1970s. Gilbert, who studies turtles, sees something similar: “Turtle biologists joke that one Tuesday in the late Tri-

assic there weren’t any turtles, and by the weekend the world was full of turtles. One reason why might be that it’s not all that hard to make a shell — all the genes are probably there already, and it doesn’t take many changes to get a shell.”

Stuart Newman, a developmental biologist at New York Medical College, takes such ideas further than most, arguing that the abilities that cells have to self-organize into complex structures can lead to major evolutionary innovations such as the origin of the vertebrate limb — a problem on which he collaborates with Altenberg’s other organizer, evo-devo researcher Gerd Müller of the University of Vienna, Austria⁴ — with



“Our way of looking at evolution counters genetic determinism and its political implications.”
— Eva Jablonka



The Altenberg Sixteen at the Konrad Lorenz Institute for Evolution and Cognition Research.

perhaps little or no genetic change. “You can’t deny the force of selection in genetic evolution,” says Newman, “but in my view this is stabilizing and fine-tuning forms that originate due to other processes.”

The same process might have given rise to animals themselves. The further you turn back the clock through geological time, Newman believes, the weaker genetic regulation of development becomes relative to plasticity and self-organization. The development of the most basic features of multicellular organisms some 600 million years ago, in the late Proterozoic, might have been the rapid and spontaneous result of molecules already present on unicellular organisms doing new jobs when cells stick together⁵. “You don’t need incremental change under gradual selection regimes to get attributes such as segmented, hollow or multilayered bodies,” says Newman. “You can get it all with thermodynamics and self-organization.”

The problem is testing such ideas. Newman suggests that knocking out the genes that stabilize development in model laboratory organisms might provide insights, but extrapolating back from modern organisms to their distant ancestors is fraught with problems. It is difficult to see how such an approach can get beyond the theoretical, says Budd, adding that what evidence there is weighs against Newman’s hypothesis. “Clearly there are physical and chemical processes that affect cells,” says Budd. “But I don’t think there is any evidence at all for the idea that development was more permissive and plastic [in the Proterozoic] and that

body plans could spontaneously emerge. The fossil record shows that body plans appeared sequentially in a series of innovations, not in a misty way at the bottom of the tree.”

Confusing what can happen and what did happen is a common criticism of the ideas raised at Altenberg. For example, some lab studies lend support to the Baldwin effect: experiments with fruitflies show that following up an environmental stress with selective breeding can produce animals that show the phenotypic response to that stress without having experienced it⁶. But there is little evidence so far that genetic change in wild populations takes this course, says Wagner. “The idea that environmentally induced changes are the path-breaker for genetic fixation is an old one, but I’m not yet convinced that’s how it works in real populations,” he says.

“These notions haven’t forced us to change the neo-darwinian paradigm,” says Jerry Coyne, an evolutionary geneticist at the University of Chicago. Coyne has little time for “evo-devo-tees”⁷ who think that the discipline will cause a revolution in biology. Researchers coming at evolution from population genetics are particularly resistant to any attempt to displace natural selection from the place

at the heart of evolutionary theory that the modern synthesis provided it with. “The whole thing about natural selection being an insufficient paradigm seems grossly overblown,” says Coyne. “There are a lot of interesting new things coming out that will change our view of evolution. But to say the modern synthesis is incomplete or fatally flawed is fatuous.”

And it is worth noting that you can work in evo-devo and not subscribe to such ideas. Sean Carroll of the University of Wisconsin in Madison sees things in terms of bridge-building, not replacement. “What did population

genetics and palaeontology have to do with each other for the past 80 years? Nothing. The modern synthesis describes evolution within populations — it’s agnostic or silent about the cumulative effect of that process,” he says. By revealing the genetic basis of development, and showing how genetics relates to morphology, evo-devo “sits right in the middle” of the two disciplines, says Carroll.

The true message of evo-devo, Carroll says, is that developmental processes have evolved in a way that allows small aspects of form to be tweaked without affecting the whole organism — something which tends to reinforce



“Evo-devo was the synthesis between evolution and development — now it is what needs to be done to get there.”
— Alan Love

B. LORENZ/KLI

K. MACWILLIAMS/UNIVERSITY OF MINNESOTA

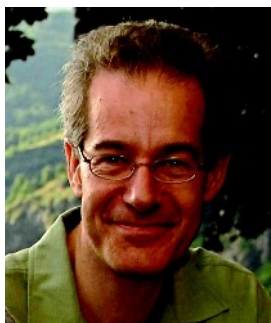
G. DUNN

the modern synthesis's view of evolution as incremental⁸. "Because we can get large effects when we manipulate genes in development, the spectre that these things have happened in history is out there," says Carroll. "But just because we can make freaky-looking animals in one step, I'm unwilling to say that evolution works that way." Wagner and his colleagues have recently shown that altering many genes in mice produces only a small effect⁹, countering the idea that most individual genes have such a wide-ranging influence that changing them would be fatal.

The differences of opinion suggest that, although evo-devo may once have looked as if it would unify population genetics and development, so far it has done more to give new voice to important problems that had been pushed to the margin — this was a strong note at Altenberg, making the meeting as much about revivalism as revolution. "Originally, the idea was that evo-devo was going to be the synthesis between evolution and development — now it is part of what needs to be done to get there," says Alan Love, a philosopher of science at the University of Minnesota in Minneapolis who attended Altenberg. "There is still a lot of outstanding work to do on fitting the pieces together, but no consensus on how to go about that right now." Nevertheless, he says, that's no cause for alarm. "What is needed is to incorporate empirical findings into the bigger picture. It took populations genetics 25 years to do that and make the modern synthesis. As far as evo-devo goes, I'd say we're smack dab in the middle of that process."

Preaching to the converted?

David Krakauer, an evolutionary theorist at the Santa Fe Institute in New Mexico who was not at Altenberg, agrees. "It's a matter of finally unifying two areas that haven't spoken to one another," he says. "To tackle any modern problem in evolutionary biology, you'll have to use development and the dynamics of the genes that underlie it." He's quite enthusiastic about the possibility of bringing together mathematical theories of pattern formation, of the kind favoured by Newman, and the large body of theory on genetic change between generations used by population geneticists such as Coyne. But at the same time, he can see forces beyond the content of the theories that may keep them apart: "It's not about totally incompatible world



"If there's one thing we don't want, it's that there's a bunch of theories out there, and they're all equal."

— Massimo Pigliucci

views, it's about who holds the torch — who are the legitimate heirs to the Darwinian intellectual estate."

Love saw the Altenberg meeting as an attempt to bridge the divide, but one that, by avoiding conflict (partly through invitations being declined), ended up a little one-sided. "Altenberg was an attempt to pull people together; the hard part was that it didn't pull in people who were less than sympathetic towards one another," he says. "It could have been a much more eraser-throwing meeting, but there is no reward for organizing that — you don't get another grant by trying to get people in the same room, you just have to take time away from the lab or fieldwork."

And there are forces at play beyond jockeying for disciplinary prestige. Never mind what can happen and what did happen. What should happen? It's a fight that evolutionary theory — rooted as it is in a world view shaped by Victorian capitalism — has always found itself dragged into. To give one example, the championing of 'punctuated equilibrium' in the fossil record by Gould and Eldredge was easily construed by participants on both sides of the debate in the 1970s as an attack from the political left — part of a broader rising of hackles at the arrival of sociobiology, selfish genes and the like. Evolutionary ideas and political metaphors still seem to seek each other out — in an extended synthesis, says Gilbert, "the gene will be a much more constitutional monarch, taking instructions from the cell and environment".

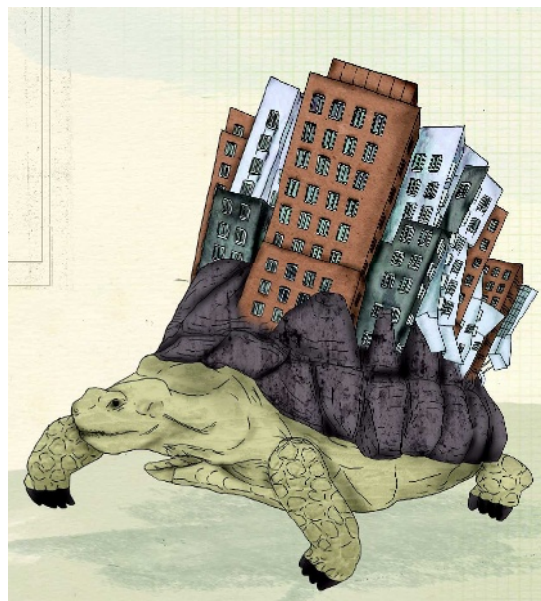
Eva Jablonka of the University of Tel Aviv, Israel, is explicit about a political side to her work. She advocates the importance of epigenetic inheritance — traits that can be passed on without changes to DNA sequence. These can be induced by environmental stressors such as temperature, diet or environmental chemicals. Such mechanisms, and insults, may be behind some inherited diseases, she says, in which case we have a responsibility to curb and reverse them. "There are social implications to our approach," says Jablonka. "Our way of looking at heredity and evolution counters genetic determinism and its political implications." Jablonka is one of

the Altenberg attendees most comfortable with the term 'postmodern'.

Yet there was no sense at Altenberg of a desire to attack evolutionary theory from the left. Quite the reverse — the dominant political concern was a fear of attack from fundamentalists. As Gould discovered, creationists seize on any hint of splits in evolutionary theory or dissatisfaction with Darwinism. In the past couple of decades, everyone has become keenly aware of this, regardless of their satisfaction or otherwise with the modern synthesis. "You always feel like you're trying to cover your rear," says Love. "If you criticize, it's like handing ammunition to these folks." So don't criticize in a grandstanding way, says Coyne: "People shouldn't suppress their differences to placate creationists, but to suggest that neo-Darwinism has reached some kind of crisis point plays into creationists' hands," he says. It is tempting to say that it's not just genes that express themselves in an environment that responds and reshapes itself around them, feeding back and complicating matters beyond simple cause and effect; the same applies to ideas. And if that seems a bit self-referential — well, that's postmodernism. ■

John Whitfield is the author of *In the Beat of a Heart: Life, Energy, and the Unity of Nature*.

1. Pigliucci, M. *Evolution* **61**, 2743–2749 (2007).
2. Gerhart, J. & Kirschner, M. *Proc. Natl Acad. Sci. USA* **104**, 8582–8589 (2007).
3. Abzhanov, A., Protas, M., Grant, B. R., Grant, P. R. & Tabin, C. J. *Science* **305**, 1462–1465 (2004).
4. Newman, S. A. & Muller, G. B. *J. Exp. Zool.* **304B**, 593–609 (2005).
5. Newman, S. A. & Bhat, R. *Phys. Biol.* **5**, 015008 (2008).
6. Braendle, C. & Flatt, T. *BioEssays* **28**, 868–873 (2006).
7. Hoekstra, H. E. & Coyne, J. A. *Evolution* **61**, 995–1016 (2007).
8. Carroll, S. B. *Cell* **134**, 25–36 (2008).
9. Wagner, G. P. et al. *Nature* **452**, 470–472 (2008).



CORRESPONDENCE

More ground work needed to prepare students for PhDs

SIR — John Kirkland's review of *Toward a Global PhD?* by M. Nerad and M. Heggelund (*Nature* **454**, 408; 2008) emphasizes the increasing popularity of doctoral degrees among students around the world. As a graduate students' representative, I have noticed a general and worrisome side effect of this trend in the three universities on two continents where I have been based.

It is striking how many early graduate students around the world complain about being unprepared for science itself, despite having passed their undergraduate degrees with ease. For graduate ecologists, for instance, it is crucial that the experimental design and statistics should be carefully thought through before the project is started. However, despite most universities' requiring a basic course in statistics, many graduates struggle to create properly designed experiments.

Ways around this problem could include practical experience on a project-based MSc course as a prerequisite for enrolling for a PhD, stemming the trend to shorten PhD courses and enhancing the quality of supervision.

Ideally, students should be motivated to learn for themselves and should have access to good library facilities so that they can acquire all the information they need before beginning data collection. But some will struggle to make the leap from an initial grasp of theory to the complexities of experimental design and analysis, and will suffer from this for the remainder of their PhD.

This large gap between undergraduate and graduate courses inevitably leads to money being wasted on badly designed projects that will never be published, and the demoralization

of graduates who will never go on to become scientists.

Cristina Banks-Leite Departamento de Ecologia, Universidade de São Paulo, 05508-900, Brazil and Division of Biology, Imperial College London, Ascot, Berkshire SL5 7PY, UK
e-mail: banks@ib.usp.br

Talking about the creative use of enhancements

SIR — In his review 'Inside the mind of a marathon runner' (*Nature* **454**, 583–584; 2008), Andy Miah says he approaches the book *What I Talk About When I Talk About Running* from the perspective of the ethics of biotechnical enhancements. But the striking title of Haruki Murakami's work deserves further comment.

The title echoes that of one of the most important short-story collections in contemporary American literature, taken from one of the stories it contains: Raymond Carver's *What We Talk About When We Talk About Love* (Knopf, 1981). That story's themes include death, isolation, drinking and different types of love. Murakami's substitution of "running" for "love" speaks for itself, as does the insularity evoked by moving from "we" to "I".

This shift towards isolation is also picked up in Alan Sillitoe's poignant *The Loneliness of the Long-Distance Runner* (W. H. Allen, 1959): a phrase that has become better known than the powerful short story that evoked it.

Carver himself even fits into Miah's theme of biotechnical enhancements. Some say he did his best work while drinking and smoking, which may have helped ease him into his writing in much the same way as Murakami's coating of Vaseline helped him slip into his triathlon wetsuit.

The term 'enhancement' could also be applied to the work of the aggressive editor Gordon Lish, who in many ways shaped Carver's unique voice. It was Lish

who coined the reflective 'What We Talk About When We Talk About Love', an improvement over the story's original and less evocative title 'Beginners'.

Whether we are talking about writing, sport or science, some enhancements pay off.

Jennifer Manning American Committee for the Weizmann Institute of Science, 633 Third Avenue, New York, New York 10017, USA
e-mail: jennifer@acwis.org

European grants: a different view puts rich countries ahead

SIR — Another look at the data used by F. Frischknecht in his Correspondence 'Small countries are unexpected winners in ERC grant tables' (*Nature* **454**, 690; 2008) reveals a different story.

Thirty-eight countries are eligible for European Research Council (ERC) grants. Of the 17 that did not receive any grants, 13 had a population size below the median (based on purchasing power parity, using data from the CIA world factbook at <https://www.cia.gov/library/publications/the-world-factbook/index.html>). At a glance, this might suggest that the smaller countries are the losers.

The reason why this is untrue is obvious, but Frischknecht's figures contain a similar error. He concludes that small countries were successful, because he omitted the small countries that did not receive any grants. Of course, larger countries are less likely to have no successful applicants, because they draw from a larger pool of talent. Consequently, Frischknecht's results are automatically biased.

If we look at the full data and use a log-linear model to regress the number of successful applicants against log-transformed population size and gross domestic product (GDP) per person, we find that the population effect (coefficient of 0.9, standard error 0.11) is not

distinguishable from 1 ($P=0.3$), indicating no evidence of bias towards smaller countries.

The wealth effect is much more important: doubling GDP per head increases the number of grants tenfold (coefficient of 2.3, standard error 0.49). One might conclude that the winners are the rich countries. Again.

Bob O'Hara Department of Mathematics and Statistics, PO Box 68 (Gustaf Hållströmin katu 2b), FIN-00014 University of Helsinki, Finland
e-mail: bob.ohara@helsinki.fi

European grants: a lifeline in poorly funded countries

SIR — I was amused to read in Friedrich Frischknecht's Correspondence 'Small countries are unexpected winners in ERC grant tables' (*Nature* **454**, 690; 2008) that Israel is ranked first in grants received, in relation both to population and to gross domestic product. This is shown in his table analysing grant awards from the European Research Council (ERC) for young-investigator research by country. No doubt this ranking will be used for political capital by various interested parties.

Although this success rate must reflect the excellence of our young scientists, I suspect that it also reflects their degree of desperation. In a country with an average local grant size of US\$30,000–40,000 and a finance ministry that is indifferent or actively opposed to funding academic science, an ERC grant is not just a glittering prize, it is one of the few available options for scientific survival. Rather than using Frischknecht's table as a measure of quality, I submit that its primary utility could be in highlighting countries that do not provide adequate local funding for their scientists.

Mike Fainzilber Department of Biological Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel
e-mail: mike.fainzilber@weizmann.ac.il

BOOKS & ARTS

Time to can the tuna fisheries?

A tale of the endangered Atlantic bluefin tuna reveals how overfishing and unreported catches threaten to wipe out this majestic predator. But it stops short of offering original solutions, explains **Geoff Arnold**.

Tuna: A Love Story

by Richard Ellis

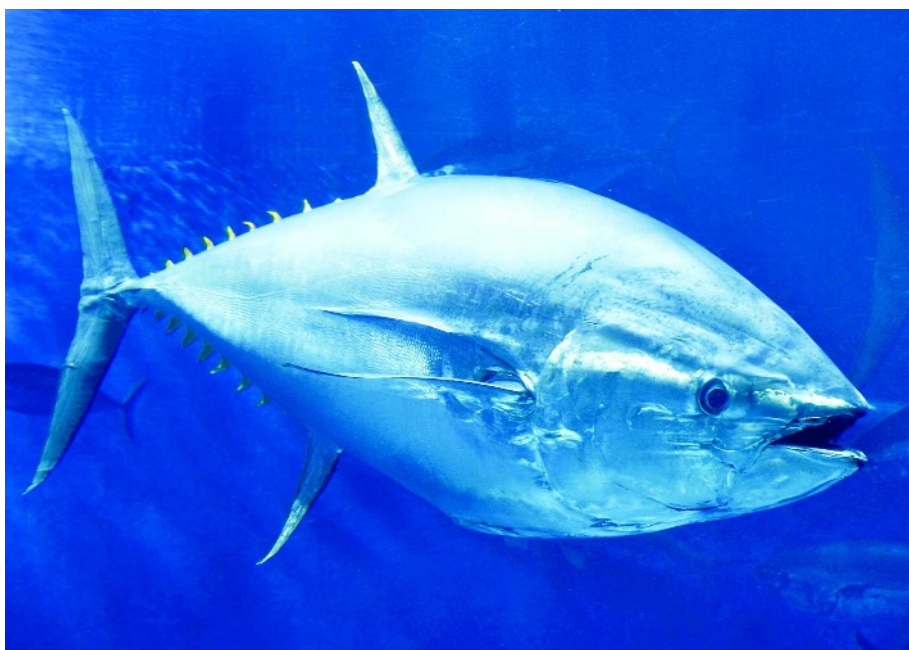
Knopf, 2008. 352 pp. \$27.95

Stocks of bluefin tuna are at risk of commercial and biological collapse worldwide, particularly in the Mediterranean Sea. Scientists are laying the foundations for better long-term management of tuna stocks by including electronic tag data, genetics and microchemistry in their assessment models. Despite their efforts, effective management action could come too late to save this majestic fish.

In his book *Tuna*, Richard Ellis makes an impassioned plea to rescue the bluefin before it is fished to extinction. *Tuna* lacks the discipline of Mark Kurlansky's *Cod* (Jonathan Cape, 1998) but Ellis conveys well the enthusiasm people have for these superbly evolved top predators that are a source of great wealth as well as prized sushi and sashimi.

Much can be learned about fish behaviour from electronic tagging. When I was involved in tracking individual plaice and cod using sonar in the North Sea in the 1970s, the process was arduous and slow. Acoustic tags could transmit for several days, but following them with a small research vessel meant bad weather often stopped the work first. To overcome these constraints we developed small data loggers, which recorded fish depth and sea temperature every few minutes without the need for tracking. Fishermen returned recaptured tags for a small reward and the data allowed us to deduce the vertical and horizontal movements of groups of fish over many months. Some 30 years on, new sensors under trial will soon allow feeding and spawning to be studied in free-ranging fish at sea, rapidly filling important gaps in biological knowledge.

Evolving tag technology has made a major contribution to tuna biology. Thirty years ago, Frank Carey built large acoustic tags at the Woods Hole Oceanographic Institution in Massachusetts, and later used them to demonstrate that tuna, swordfish and some sharks can regulate the temperature of some parts of their bodies. Recent electronic tagging programmes have revealed transoceanic migrations of bluefin tuna in the Atlantic, Pacific and Southern Oceans, and shown the widespread dispersion and overlap of Atlantic bluefin populations that



R. WILDER/MONTEREY BAY AQUARIUM

Calls are being made for a global ban on bluefin tuna fishing before it is too late.

segregate to discrete spawning grounds in the Mediterranean and Gulf of Mexico.

Much of this story can be gleaned from Ellis's book, which deals mainly with bluefin tuna but includes yellowfin, skipjack, albacore and bigeye tuna. The chapter about sport fishing, describing millionaires such as popular author Zane Grey who pursued giant bluefin with a rod and line from small boats, is exciting and enjoyable. Other chapters dealing with biology, systematics, commercial fisheries, marketing and ranching are less well focused. Material is distributed haphazardly with much repetition, some sections of text are repeated verbatim and the use of references is erratic. Catch statistics and quotas pop up everywhere and at least six chapters discuss assessment and management of tuna stocks without synthesis. Although the footnotes are models of clarity, the main text is occasionally pretentious, and the lack of a glossary makes it difficult to resolve apparent contradictions in the use of some technical terms. Happily, the book is well illustrated and Ellis's line drawings are excellent.

Tuna will leave readers appalled that Atlantic and Southern bluefin stocks are so close to collapse and that deliberate underreporting of catches by key fishing nations has been a major factor in the tuna's decline. However,

the book does not advance any original solutions and the logic of its concluding chapters is muddled. Ellis uses biblical sources to justify conservation and dismisses fisheries management in favour of farming, even though it has been impossible to get bluefin to spawn in captivity until very recently. It would have been more constructive had Ellis weighed the relative merits of both approaches, reiterated the continuing need for strong conservation of wild populations and summarized what can be done.

In March this year, chastened by criticism of its failure to manage the bluefin tuna fishery in the Mediterranean and eastern Atlantic, the European Commission instituted a strong monitoring and enforcement plan to prevent a recurrence of the substantial overfishing of bluefin recorded by European Union member states in 2007 (<http://tinyurl.com/6orrrz>). By mid-June, when the 2008 quota of 28,500 tonnes was exhausted, the European Commission called a halt to purse seine fishing, which uses the huge drawstring nets responsible for more than 70% of the total tuna catch. It also prohibited transshipment and transfer of fish to fattening cages from either EU or non-EU vessels.

Regrettably, the EC's apparently strong measures will almost certainly fail to achieve

stock recovery unless the International Commission for the Conservation of Atlantic Tunas (ICCAT), which sets the quotas, reforms the ineffective 15-year recovery plan it adopted in 2006.

When it meets in November this year, the ICCAT must adopt the measures advocated by its own Standing Committee on Research and Statistics (<http://iccat.int/scrs.htm>). These include closure of the Mediterranean throughout the spawning season, full enforcement of the increased minimum size for caught tuna, and a catch limit of 15,000 tonnes — less than a third of recent totals — for the next few years. It should also heed the call of the conservation group WWF, which is asking for a

three-year moratorium on bluefin tuna fishing (<http://tinyurl.com/5ehxtw>). With far too many fishing vessels, significant underreporting of catches by several EU countries and considerable illegal, unregulated and unreported fishing by neighbouring states such as Turkey, Croatia and Libya, time is fast running out for the Mediterranean bluefin tuna. Its loss would be an economic and ecological disaster. ■

Geoff Arnold is former Chair of the Fisheries Science and Management Group at the Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft, UK, and a member of the SCOR Panel on New Technologies for Observing Marine Life.

e-mail: geoff@arnold45.freemove.co.uk

2001 discovery of a strange squid with spaghetti-like arms. A news kiosk features frequent ocean updates. Exhibit developer Jill Johnson hopes the displays will educate visitors and encourage them to be better stewards of the world's waters.

Looming over the hall is a full-scale model of a female North Atlantic right whale called Phoenix. Scientists have spent around 20 years tracking the real Phoenix, distinguishable by her unique scars. The 14-metre-long mother of three was recently spotted in the Gulf of Maine, and is one of an estimated 350 North Atlantic right whales left in the ocean.

One of the strengths of the Smithsonian is its vast collection of marine species, which the hall's curators were eager to display. Specimens are traditionally stored in alcohol, but a change in the fire regulations in 2005 meant that fire-safety experts would allow only a maximum of 38 litres of alcohol in the 2,000-metre-square space, explains Johnson.

The answer came from technology company 3M, based in St Paul, Minnesota. One of their non-flammable, non-toxic hydrofluoroether fluids, developed as a cleaning fluid and solvent, forms an envelope around samples to seal in the formalin preservative. 3M donated some 4,000 litres of the stuff to the museum, where it now surrounds various artefacts, including a coelacanth and her pup, a prehistoric fish that was thought to be extinct until its rediscovery in 1938.

The hall's specimens are the first to be displayed in the fluid, which, unlike alcohol, does not leach colour. "A lot of other museums are really interested in how this works out," Vecchione says.

One of the hall's stars is a seven-metre-long giant squid that rests majestically in a coffin-like tank filled with the liquid. Along with a three-metre-long male squid, it is on loan from the Spanish research organization CEPESMA. Transporting the pair across the Atlantic in 1,500 litres of formalin was quite a task. The US Air Force and Navy helped out with the trip, dubbed 'Operation Calamari'. Elizabeth Musteen, who coordinated the journey, was concerned about the weight of the cargo until an unimpressed Air Force representative chastised her: "Ma'am, I move tanks. I can do this."

Also on show is a hand-carved Tlingit canoe, some 8 metres long. The Pacific northwest tribe is one of many cultures represented in the hall that depend on the sea. As Johnson explains, "The main message is that the ocean, singular, is a global system that is essential to all life — including yours". ■

Amber Dance is a News intern based in *Nature's* Washington DC office.

Ocean exhibit makes waves

The Sant Ocean Hall

Smithsonian National Museum of Natural History, Washington DC

Opens on 27 September 2008

Geographers will tell you that there are four world oceans, or five if you count Antarctic waters. The designers of the new Sant Ocean Hall at the Smithsonian National Museum of Natural History, based in Washington DC, beg to differ. Although it acknowledges distinct ocean basins, the exhibit emphasizes that the world's waters are linked in a single ecosystem that influences all life on Earth.

The US\$49-million hall is due to open on 27 September and will house the museum's largest permanent display. A joint venture with the US National Oceanic and Atmospheric Administration, the hall explains oceanography, biology, ecology and anthropology.

It was inspired by aspects of the ocean in crisis, exploring the effects of oil spills, climate change and ocean acidification. Other sections highlight the ocean's bounty and the resources it provides.

Changing displays showcase scientific research, including the Census of Marine Life — a collaborative project to assess marine diversity — and curator Michael Vecchione's



An artist's impression of the Sant Ocean Hall shows its emphasis on the oceans as a single ecosystem.

stock recovery unless the International Commission for the Conservation of Atlantic Tunas (ICCAT), which sets the quotas, reforms the ineffective 15-year recovery plan it adopted in 2006.

When it meets in November this year, the ICCAT must adopt the measures advocated by its own Standing Committee on Research and Statistics (<http://iccat.int/scrs.htm>). These include closure of the Mediterranean throughout the spawning season, full enforcement of the increased minimum size for caught tuna, and a catch limit of 15,000 tonnes — less than a third of recent totals — for the next few years. It should also heed the call of the conservation group WWF, which is asking for a

three-year moratorium on bluefin tuna fishing (<http://tinyurl.com/5ehxtw>). With far too many fishing vessels, significant underreporting of catches by several EU countries and considerable illegal, unregulated and unreported fishing by neighbouring states such as Turkey, Croatia and Libya, time is fast running out for the Mediterranean bluefin tuna. Its loss would be an economic and ecological disaster. ■

Geoff Arnold is former Chair of the Fisheries Science and Management Group at the Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft, UK, and a member of the SCOR Panel on New Technologies for Observing Marine Life.

e-mail: geoff@arnold45.freemove.co.uk

2001 discovery of a strange squid with spaghetti-like arms. A news kiosk features frequent ocean updates. Exhibit developer Jill Johnson hopes the displays will educate visitors and encourage them to be better stewards of the world's waters.

Looming over the hall is a full-scale model of a female North Atlantic right whale called Phoenix. Scientists have spent around 20 years tracking the real Phoenix, distinguishable by her unique scars. The 14-metre-long mother of three was recently spotted in the Gulf of Maine, and is one of an estimated 350 North Atlantic right whales left in the ocean.

One of the strengths of the Smithsonian is its vast collection of marine species, which the hall's curators were eager to display. Specimens are traditionally stored in alcohol, but a change in the fire regulations in 2005 meant that fire-safety experts would allow only a maximum of 38 litres of alcohol in the 2,000-metre-square space, explains Johnson.

The answer came from technology company 3M, based in St Paul, Minnesota. One of their non-flammable, non-toxic hydrofluoroether fluids, developed as a cleaning fluid and solvent, forms an envelope around samples to seal in the formalin preservative. 3M donated some 4,000 litres of the stuff to the museum, where it now surrounds various artefacts, including a coelacanth and her pup, a prehistoric fish that was thought to be extinct until its rediscovery in 1938.

The hall's specimens are the first to be displayed in the fluid, which, unlike alcohol, does not leach colour. "A lot of other museums are really interested in how this works out," Vecchione says.

One of the hall's stars is a seven-metre-long giant squid that rests majestically in a coffin-like tank filled with the liquid. Along with a three-metre-long male squid, it is on loan from the Spanish research organization CEPESMA. Transporting the pair across the Atlantic in 1,500 litres of formalin was quite a task. The US Air Force and Navy helped out with the trip, dubbed 'Operation Calamari'. Elizabeth Musteen, who coordinated the journey, was concerned about the weight of the cargo until an unimpressed Air Force representative chastised her: "Ma'am, I move tanks. I can do this."

Also on show is a hand-carved Tlingit canoe, some 8 metres long. The Pacific northwest tribe is one of many cultures represented in the hall that depend on the sea. As Johnson explains, "The main message is that the ocean, singular, is a global system that is essential to all life — including yours". ■

Amber Dance is a News intern based in *Nature's* Washington DC office.

Ocean exhibit makes waves

The Sant Ocean Hall

Smithsonian National Museum of Natural History, Washington DC

Opens on 27 September 2008

Geographers will tell you that there are four world oceans, or five if you count Antarctic waters. The designers of the new Sant Ocean Hall at the Smithsonian National Museum of Natural History, based in Washington DC, beg to differ. Although it acknowledges distinct ocean basins, the exhibit emphasizes that the world's waters are linked in a single ecosystem that influences all life on Earth.

The US\$49-million hall is due to open on 27 September and will house the museum's largest permanent display. A joint venture with the US National Oceanic and Atmospheric Administration, the hall explains oceanography, biology, ecology and anthropology.

It was inspired by aspects of the ocean in crisis, exploring the effects of oil spills, climate change and ocean acidification. Other sections highlight the ocean's bounty and the resources it provides.

Changing displays showcase scientific research, including the Census of Marine Life — a collaborative project to assess marine diversity — and curator Michael Vecchione's



An artist's impression of the Sant Ocean Hall shows its emphasis on the oceans as a single ecosystem.

NATL MUS. NAT. HIST./SMITHSONIAN INST.

Stages of uncertainty

Stage Directions: Writing on Theatre 1970–2008

by Michael Frayn

Faber and Faber: 2008. 288 pp. £20

In Michael Frayn's 1998 play *Copenhagen*, physicists Niels Bohr and Werner Heisenberg, along with Bohr's wife Margrethe, reunite in the afterworld to work out the mystery of why Heisenberg visited Bohr in Copenhagen in 1941. The play has stimulated tremendous debate about the people and events it depicts. Frayn's new book includes his writings about the play both before it was staged and in response to this debate. It is useful to have them collected here alongside Frayn's other writings about theatre over the past three decades.

Stage Directions reveals much about the mind behind *Copenhagen*, a play he admits gave the 'kiss of life' to a playwriting career that was flagging in the 1990s. Just a few years earlier, in 1985, Frayn's career had been at its peak, with three of his works running in London simultaneously. One chapter, entitled 'On the Roller Coaster', charts the frenzied, out-of-control lurching between failure and success that characterized his early writing. A sense of precariousness has remained with him, even after unqualified stage

successes like *Copenhagen*. Frayn isn't afraid to wallow in his failures. This, together with his readiness to place the contributions of the directors and actors he has worked with far above his own efforts, makes for appealing reading.

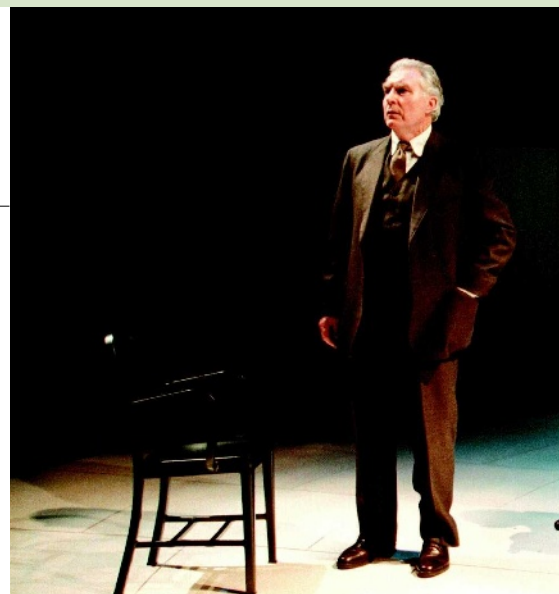
Frayn's theatre work spans everything from comedy and farce, opera and film, to serious drama and a substantial body of translations of Anton Chekhov's plays. He fondly recalls his childhood theatrical efforts, first at toy theatre and then at magic and conjuring tricks. Some hilarious mishaps, with the scenery crashing down around one particularly spectacular failure, planted the very productive seeds of his later farces such as *Donkeys' Years* and *Noises Off*.

To date, *Copenhagen* is Frayn's only venture into science on stage. Indeed, *Copenhagen* and *Noises Off* seem so remote from one another that it is hard to believe they are by the same playwright. *Copenhagen* plumbs some of the most difficult scientific and moral questions of our time; in a sense, it forms a trilogy with Frayn's two most recent plays, *Democracy* and *Afterlife*. Apart from the single-word titles, these all reflect Frayn's abiding Germanic interests and fascination with 'great men' on the stage of history — Heisenberg and other physicists in *Copenhagen*, German chancellor Willy Brandt

in *Democracy* and the Austrian (later American) theatre director Max Reinhardt in *Afterlife*. These men all met with failure and success. The plays probe the extent to which personal characteristics brought about these failures, and to what degree they were caused by political circumstances that overwhelmed the characters.

All three plays reflect Frayn's interest in 'the epistemology of intention', or the question of how we attempt — and fail — to know our own and others' thoughts and motives. The book reveals that this theme is already latent in his earliest plays. Another, more overt theme that he develops in his work is how we impose our ideas on the world around us.

Because *Copenhagen*, *Democracy* and *Afterlife* all involve real historical figures and situations,



D. COOPER/PHOTOSTAGE

Saving plastics for posterity

Conservation of Plastics: Materials Science, Degradation and Preservation

by Yvonne Shashoua

Butterworth-Heinemann: 2008. 300 pp. \$87.95, £49.99

Many modern cultural artefacts, from Neil Armstrong's space suit to works of art, contain plastics that may degrade readily. Museums now face the challenge of saving vulnerable and important examples for future generations. Yvonne Shashoua's book *Conservation of Plastics* is a timely, thoughtful and thorough review of the young discipline of plastics conservation. It will help conservators, curators, archivists and collectors of modern materials, and will be of interest to materials scientists and chemists.

Often misidentified by the untrained eye, plastics are present in scientific, technological and social historical collections, and increasingly in modern art and textiles. Certain twentieth-century plastics, notably early polyvinyl chloride and polyurethane foams, degrade relatively swiftly and severely. Casualties include design

icons such as the inflatable 1960s Blow Chair and highly collectable vintage Barbie dolls.

Semisynthetic plastics from the nineteenth century, especially cellulose nitrate (celluloid) and cellulose acetate, are particularly susceptible to degradation, whether the stimulus is physical, chemical or environmental. Irreversible damage to early celluloid cine film and its replacement, cellulose acetate safety film, pose a great problem for archivists. The earliest films and the historical evidence they contain may be lost unless they are transferred onto more robust media. The twentieth-century sculptor Naum Gabo used cellulose nitrate and cellulose acetate in his early works. Some have deteriorated so badly that conservators are turning to computer-aided methods to record their form.

Early steps to save plastics were discussed at a 1988 conference held by the Scottish Society for Conservation and Restoration. Several books followed. The 1993 Canadian Conservation Institute publication *Saving the Twentieth Century*, edited by David Grattan, established the discipline of plastics conservation. In their 1999

book *Plastics: Collecting and Conserving*, Anita Quye and Colin Williamson issued guidelines for the care of plastics in collections to acceptable museum and conservation standards.

A decade on, Shashoua's welcome and accessible book provides up-to-date, professional advice on caring for plastics, including associated risks and useful data. She provides context with introductory chapters on plastics in collections and the technology of plastics and their historical development, and uses her background as a polymer chemist and conservator to explain the nature of plastics, their types and related chemistry. Especially helpful are tables of frequently used additives, properties of plastics and the impact of chemical reactions in plastics, notably those caused by light, heat and oxygen.

To assist with the difficulty of identifying plastics, the book lists characteristic odours, recycling symbols and specific gravities. Shashoua discusses identification tests, including the controversial 'hot pin test', in which glowing hot pins are applied to the surface of a plastic object and note taken of the resulting marks and odours. This method is favoured by collectors and antique sellers, but rejected by museum professionals owing to the damage



Michael Frayn's fascination with great men of history prompted his play *Copenhagen*.

Frayn provides substantial 'postscripts' to each one, reprinted in the book, in which he lays out his source material, explains how he used it, and where and why he diverges from the historical record. For *Copenhagen*, he retraces the tremendous research he did into the physics and the physicists, from concepts of uncertainty and complementarity in quantum mechanics, to calculations of the critical mass of fissile material needed to make an atomic bomb, and the extent and nature of the German atomic weapons programme. His assessments of Bohr and Heisenberg are detailed and perceptive.

Frayn also discusses the 1945 detainment of Heisenberg and other German physicists at Farm Hall, a manor house outside Cambridge, UK, noting that 'the story of Farm Hall

is another complete play in itself'. If only he would one day write it.

In a post-postscript, Frayn recounts how *Copenhagen* took on a life of its own and even changed the historical record, prompting first the revelation of draft letters from Bohr to Heisenberg about their 1941 meeting, and then the discovery of a letter Heisenberg wrote to his wife during that meeting, which took place, it turns out, on three separate days over a week. Although he wishes he had gained access to some of the documents earlier, especially Bohr's draft letters, he maintains that in the end, the play's integrity still stands.

One can see why he is drawn to Chekhov. The Russian, like Frayn, wrote novels and short stories before coming to playwriting late in life, for which he is arguably best known. Frayn's sympathy with Chekhov's 'struggles to understand and master the recalcitrant medium of the theatre' comes through in this fascinating collection: how difficult theatre is, by turns unpredictable, maddening, tedious, enthralling, crushing, exhilarating. In other words, just like science. ■

Kirsten E. Shepherd-Barr is a lecturer in modern drama at the University of Oxford, and a fellow and tutor in English at St Catherine's College, University of Oxford, UK. She is author of *Science on Stage: From Doctor Faustus to Copenhagen*. e-mail: kirsten.shepherd-barr@ell.ox.ac.uk

that can be incurred. She summarizes analytical techniques, including Fourier transform infrared and Raman spectroscopy, and supplies technical data that are useful to analytical chemists.

Shashoua includes dramatic case studies, such as the polyethylene head that was used to display a gas mask with vulcanized rubber strips. These had reacted chemically with the polyethylene so that the plastic head suffered irreversible and disfiguring degradation. This shows that sulphur-containing vulcanizing agents diffused from the gas mask into the polyethylene, and highlights the importance of using inert materials as display mounts for plastics.

Shashoua outlines available treatments, pitfalls associated with using solvents and the risks of cleaning and labelling plastics. She focuses on objects rather than media such as film, which is a specialization in itself. She highlights a need for better training for plastics curators, collectors and conservators, and outlines a strategy for surveying collections of plastics for degradation.

Global plastics production is rapidly increasing — around 230 million tonnes were produced in 2007. The compositions of emerging specialist plastics, such as composites, copolymers and

blends, will make them even more difficult to identify and conserve and will challenge future collectors. With limited museum resources available, the future of plastics conservation will be subject to strict cost-benefit analysis. Saving Armstrong's space suit, for example, may warrant more expenditure than rescuing everyday, mass-produced plastic goods.

The book also touches on the conflicts between traditional conservation ethics and their application to new, short-lived materials, such as the biodegradable plastics being developed to offset mounting concerns about waste in landfills, landscapes and oceans. Biodegradable plastics provide increasing challenges to conservators as different types are being produced for an ever wider range of applications, from packaging to mobile phone casings.

The conservation of plastics is maturing as a discipline, but challenges lie ahead. Shashoua's guidelines in *Conservation of Plastics* do not supply all the answers, but they ensure the best chance of survival for current and future plastics collections into the twenty-second century. ■

Susan Mossman is curator of *Plasticity: 100 Years of Making Plastics* at The Science Museum, London SW7 2DD, UK, and author of *Fantastic Plastic*.

Robots refined

The fifth ArtBots exhibition, running at Trinity College Dublin from 19–21 September, aims to challenge the traditional view of robots by showing that they can be art.

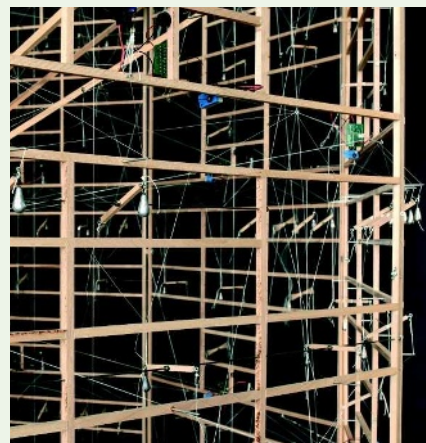
Just 15 of the 100 robots submitted have made this year's cut. Pete Redmond's *RuBot II* comes closest to what most expect a robot to be. With its humanoid torso, webcam eyes and pneumatic arms, the robot can solve a Rubik's cube in an average of 35 seconds.

Many of the works displayed subvert such stark efficiency. Allison Kudla's fragile-looking *The Search For Luminosity* frames six *Oxalis* plants with high-tech sensors and lights. When the plants raise their leaves looking for the expected Sun, a system of sensors turns on the exhibit's lights to provide sustenance. "The idea is about questioning and dismantling the hierarchy between the biological and the physical," says Kudla.

Other robots in the show look like they might date from before the Industrial Revolution. The most striking example is Ralf Baecker's *Rechnender Raum* ('Calculating Space'; pictured). "All the logic gates like and/or/not are formed by levers, strings and weights," explains Baecker. These units are connected by strings that communicate digital signals: a pulled string equals 1, a loose string equals 0. Through these strings, each cell in the framework receives and transmits instructions to neighbouring cells, resulting in a huge, primitive computer that sits and calculates with no regard to the viewer.

The show also features a panel discussion titled 'Are we living in a robotic cargo cult?'. Participants will discuss how ArtBot projects can confront the issue of whether current and future robots are "shaped by unrealistic and potentially de-humanizing science fiction fantasies". Their conclusions will feed in to an EU-funded project, *Living with Robots and Interactive Companions*. ■

Daniel Cressey is a reporter for *Nature*.





Michael Frayn's fascination with great men of history prompted his play *Copenhagen*.

Frayn provides substantial 'postscripts' to each one, reprinted in the book, in which he lays out his source material, explains how he used it, and where and why he diverges from the historical record. For *Copenhagen*, he retraces the tremendous research he did into the physics and the physicists, from concepts of uncertainty and complementarity in quantum mechanics, to calculations of the critical mass of fissile material needed to make an atomic bomb, and the extent and nature of the German atomic weapons programme. His assessments of Bohr and Heisenberg are detailed and perceptive.

Frayn also discusses the 1945 detainment of Heisenberg and other German physicists at Farm Hall, a manor house outside Cambridge, UK, noting that 'the story of Farm Hall

is another complete play in itself'. If only he would one day write it.

In a post-postscript, Frayn recounts how *Copenhagen* took on a life of its own and even changed the historical record, prompting first the revelation of draft letters from Bohr to Heisenberg about their 1941 meeting, and then the discovery of a letter Heisenberg wrote to his wife during that meeting, which took place, it turns out, on three separate days over a week. Although he wishes he had gained access to some of the documents earlier, especially Bohr's draft letters, he maintains that in the end, the play's integrity still stands.

One can see why he is drawn to Chekhov. The Russian, like Frayn, wrote novels and short stories before coming to playwriting late in life, for which he is arguably best known. Frayn's sympathy with Chekhov's 'struggles to understand and master the recalcitrant medium of the theatre' comes through in this fascinating collection: how difficult theatre is, by turns unpredictable, maddening, tedious, enthralling, crushing, exhilarating. In other words, just like science. ■

Kirsten E. Shepherd-Barr is a lecturer in modern drama at the University of Oxford, and a fellow and tutor in English at St Catherine's College, University of Oxford, UK. She is author of *Science on Stage: From Doctor Faustus to Copenhagen*. e-mail: kirsten.shepherd-barr@ell.ox.ac.uk

that can be incurred. She summarizes analytical techniques, including Fourier transform infrared and Raman spectroscopy, and supplies technical data that are useful to analytical chemists.

Shashoua includes dramatic case studies, such as the polyethylene head that was used to display a gas mask with vulcanized rubber strips. These had reacted chemically with the polyethylene so that the plastic head suffered irreversible and disfiguring degradation. This shows that sulphur-containing vulcanizing agents diffused from the gas mask into the polyethylene, and highlights the importance of using inert materials as display mounts for plastics.

Shashoua outlines available treatments, pitfalls associated with using solvents and the risks of cleaning and labelling plastics. She focuses on objects rather than media such as film, which is a specialization in itself. She highlights a need for better training for plastics curators, collectors and conservators, and outlines a strategy for surveying collections of plastics for degradation.

Global plastics production is rapidly increasing — around 230 million tonnes were produced in 2007. The compositions of emerging specialist plastics, such as composites, copolymers and

blends, will make them even more difficult to identify and conserve and will challenge future collectors. With limited museum resources available, the future of plastics conservation will be subject to strict cost-benefit analysis. Saving Armstrong's space suit, for example, may warrant more expenditure than rescuing everyday, mass-produced plastic goods.

The book also touches on the conflicts between traditional conservation ethics and their application to new, short-lived materials, such as the biodegradable plastics being developed to offset mounting concerns about waste in landfills, landscapes and oceans. Biodegradable plastics provide increasing challenges to conservators as different types are being produced for an ever wider range of applications, from packaging to mobile phone casings.

The conservation of plastics is maturing as a discipline, but challenges lie ahead. Shashoua's guidelines in *Conservation of Plastics* do not supply all the answers, but they ensure the best chance of survival for current and future plastics collections into the twenty-second century. ■

Susan Mossman is curator of *Plasticity: 100 Years of Making Plastics* at The Science Museum, London SW7 2DD, UK, and author of *Fantastic Plastic*.

Robots refined

The fifth ArtBots exhibition, running at Trinity College Dublin from 19–21 September, aims to challenge the traditional view of robots by showing that they can be art.

Just 15 of the 100 robots submitted have made this year's cut. Pete Redmond's *RuBot II* comes closest to what most expect a robot to be. With its humanoid torso, webcam eyes and pneumatic arms, the robot can solve a Rubik's cube in an average of 35 seconds.

Many of the works displayed subvert such stark efficiency. Allison Kudla's fragile-looking *The Search For Luminosity* frames six *Oxalis* plants with high-tech sensors and lights. When the plants raise their leaves looking for the expected Sun, a system of sensors turns on the exhibit's lights to provide sustenance. "The idea is about questioning and dismantling the hierarchy between the biological and the physical," says Kudla.

Other robots in the show look like they might date from before the Industrial Revolution. The most striking example is Ralf Baecker's *Rechnender Raum* ('Calculating Space'; pictured). "All the logic gates like and/or/not are formed by levers, strings and weights," explains Baecker. These units are connected by strings that communicate digital signals: a pulled string equals 1, a loose string equals 0. Through these strings, each cell in the framework receives and transmits instructions to neighbouring cells, resulting in a huge, primitive computer that sits and calculates with no regard to the viewer.

The show also features a panel discussion titled 'Are we living in a robotic cargo cult?'. Participants will discuss how ArtBot projects can confront the issue of whether current and future robots are "shaped by unrealistic and potentially de-humanizing science fiction fantasies". Their conclusions will feed in to an EU-funded project, *Living with Robots and Interactive Companions*. ■

Daniel Cressey is a reporter for *Nature*.



ESSAY

Asilomar 1975: DNA modification secured

The California meeting set standards allowing geneticists to push research to its limits without endangering public health. Organizer **Paul Berg** asks if another such meeting could resolve today's controversies.

Today, the benefits of genetic engineering, and the risks and ethical dilemmas that it presents, are part of everyday public discourse, thrashed out in newspaper columns and by politicians and commentators everywhere. In the early 1970s, it was a very different picture. Scientists were only just learning how to manipulate DNA from various sources into combinations that were not known to exist naturally. Although they were confident that the new technology offered considerable opportunities, the potential health and environmental risks were unclear.

The people who sounded the alarm about this new line of experimentation were not politicians, religious groups or journalists, as one might expect: they were scientists. They called for a worldwide moratorium on the work, followed by an international conference of experts at which the nature and magnitude of the risks could be assessed. At that gathering, the International Congress on Recombinant DNA Molecules, held at the Asilomar Conference Center in Pacific Grove, California, in February 1975, it was agreed that the research should continue but under stringent guidelines. The conference marked the beginning of an exceptional era for science and for the public discussion of science policy.

Cancer-carrying bacteria

Some of the concerns about recombinant DNA experimentation stemmed from my own work with the Simian Virus 40 (SV40), which can produce tumours in rodents. My aim was to use SV40 to introduce new genes into mammalian cells. Because the DNA of SV40 can integrate into the chromosomes of infected cells, I reasoned that any 'foreign DNA' associated with it would also become part of the infected cell's genetic make-up, and consequently we might be able to study the foreign DNA's expression in mammalian cells. To test this, we inserted a segment of DNA containing three *Escherichia coli* genes responsible for the metabolism of the sugar galactose into the genome of the Simian Virus.

Several scientists feared that bacteria carrying SV40 DNA might escape and cause cancer in people infected, so we chose to defer our experiments until we could be sure that the risk was nonexistent. Most researchers, like me, acknowledged that the new technology opened extraordinary avenues for genetics

and could ultimately lead to exceptional opportunities in medicine, agriculture and industry. But we conceded that unfettered pursuit of these goals might have unforeseen and damaging consequences for human health and Earth's ecosystems.

Earlier, in mid-1974, I had led a committee that communicated those concerns to the president of the US National Academy of Sciences and published them in *Science*, *Nature* and in *Proceedings of the National Academy of Sciences*. We recommended a voluntary moratorium on certain recombinant DNA experiments that were considered potentially hazardous. The committee was particularly concerned that introduced genes could change normally innocuous microbes into cancer-causing agents or into human pathogens, resistant to antibiotics or able to produce dangerous toxins.

Scientists around the world hotly debated the wisdom of our call for caution, and the press had a field day conjuring up fantastical 'what if' scenarios. Yet the moratorium was universally observed in academic and industrial research centres. Meanwhile, the public seemed comforted by the fact that the freeze had been proposed by the very people who had helped to develop the technology.

We also proposed an international conference at which scientists and appropriate experts could assess the risks of recombinant DNA technology and devise ways of reducing them. With the backing of the National Academy of Sciences and the National Institutes of Health, I and four others — David Baltimore, Sydney Brenner, Richard Roblin and Maxine Singer — drew up the agenda for the conference. Its main aim was to consider whether to lift the voluntary moratorium and, if so, what conditions to impose to ensure that the research could proceed safely. Non-scientists were also encouraged to join in, and the approximately 140 participants included scientists, lawyers, journalists and government officials.

As might be expected, there was considerable disagreement during the conference about the existence and magnitude of the purported risks. Some biologists and public officials were certain that recombinant DNA research was flirting with disaster and that lifting the

moratorium would be a blunder. Others argued that the research was safe. Heated discussions carried on during the breaks, at meal times, over drinks and well into the small hours. I was struck by how often scientists willingly acknowledged the risks in other's experiments but not in their own. Brenner repeatedly warned of the consequences of doing nothing, predicting that such apparently self-serving behaviour would be publicly condemned and that government interference or even legislation would follow.

The turning point

What turned the debate around was the suggestion to assign a risk estimate to the different types of experiments envisaged, and to apply safety guidelines of varying stringency according to the degree of risk. This system worked on two levels. The first was physical containment, whereby the degree of risk was matched with the type of laboratory facility required. So, experiments with little or no risk could be done on an open bench; those with some risk might require laminar flow hoods; a high risk might necessitate an airlock and a laboratory under negative pressure; whereas experiments using known human pathogens would be either prohibited or restricted to specialized facilities. Brenner suggested this should be supplemented with an additional, biological level of containment to minimize the damage should engineered organisms escape into the environment. Thus

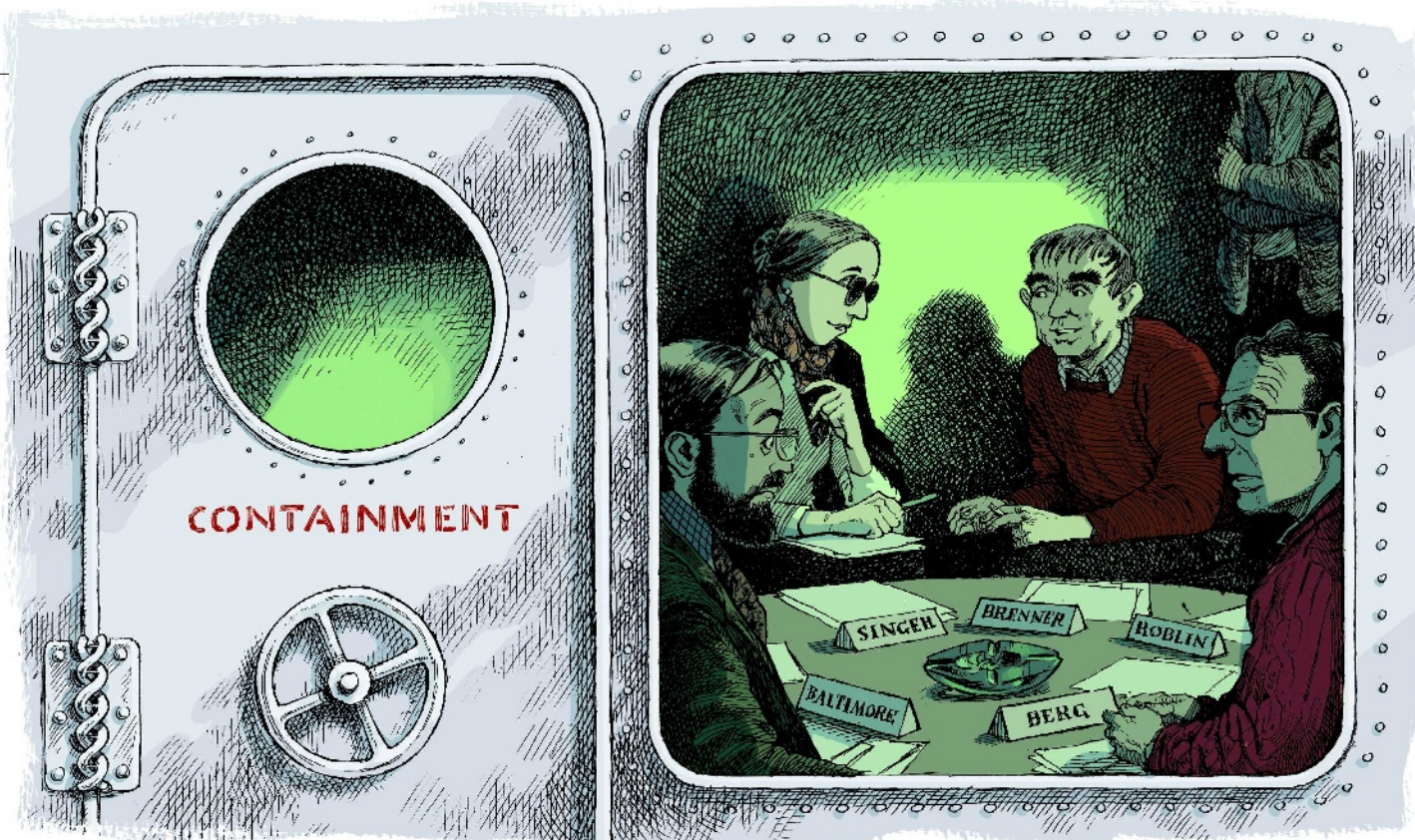
in cloning experiments that were judged to be of little or no risk, researchers could work on relatively innocuous organisms such as widely used lab strains of *E. coli* and *Bacillus subtilis*;

riskier experiments would have to use bacteria that had been genetically modified so they could not survive outside the laboratory.

Participants agreed on the final day of the conference that research should continue, but under stringent restrictions. The recommendations formed the basis of the official US guidelines on research involving recombinant DNA, issued in July 1976. They have proved remarkably effective.

In the 33 years since Asilomar, researchers around the world have carried out countless experiments with recombinant DNA without

"The press had a field day conjuring up fantastical 'what if' scenarios."



DAVID PARKINS

reported incident. Many of these experiments were inconceivable in 1975, yet as far as we know, none has been a hazard to public health. Moreover, the fear among scientists that artificially moving DNA among species would have profound effects on natural processes has substantially disappeared with the discovery that such exchanges occur in nature.

The promised land

What of the benefits of allowing researchers to continue to work with recombinant DNA? In 1975, many scientists predicted that the technology would soon yield important drugs, industrial products and improved agricultural varieties.

In fact, such developments have taken longer than anticipated. Some have never been realized because identifying the genes responsible for producing certain products or conditions, and learning how to manipulate them usefully has been more difficult than expected. Since the mid-1980s, however, the number of products has risen continually. Genetically modified hormones, vaccines, therapeutic agents and diagnostic tools are enhancing medical practice. Genetically engineered food plants are being grown and sold for consumption in both developed and developing countries. A thriving biotechnology industry has created products, jobs and wealth for scientists and others. Very few Asilomar attendees foresaw this great potential — nor could we have

predicted the pace at which our fundamental understanding of biology has grown.

Apart from laying the foundations for an effective safety regime, what else did Asilomar achieve? First and foremost, I feel that scientists were able to gain the public's trust — something that is now much more difficult for researchers working in biotechnology. Because some 15% of the participants at Asilomar were from the media, the public was well informed about the deliberations, as well as the bickering, accusations, wavering views and ultimately the consensus. Many scientists feared that a public debate would place crippling restrictions on molecular biology, but the effort encouraged responsible discussion that led to a consensus that most researchers supported.

Could an Asilomar-type conference help resolve some of the controversies now confronting scientists and the public — such as over fetal tissue, embryonic stem-cell research, somatic and germ-line gene therapy and the genetic modification of food crops? I believe that it would be much more difficult to organize such an event today. In the 1970s, most of the scientists engaged in recombinant DNA research were working in public institutions and were therefore able to get together and voice opinions without having to look over their shoulders. This is no longer the case — as many scientists now work for private companies where

commercial considerations are paramount.

Related to this is that so many issues in science and technology today are beset by economic self-interest and, increasingly, by nearly irreconcilable ethical and religious conflicts, as well as by challenges to deeply held social values. A conference that sets out to find a consensus among such contentious views would, I believe, be doomed to acrimony and policy stagnation.

That said, there is a lesson in Asilomar for all of science: the best way to respond to concerns created by emerging knowledge or early-stage technologies is for scientists from publicly-funded institutions to find common cause with the wider public about the best way to regulate — as early as possible. Once scientists from corporations begin to dominate the research enterprise, it will simply be too late. ■

Paul Berg was one of the organizers of the International Congress on Recombinant DNA Molecules held in Asilomar, 24–27 February 1975. He is Cahill professor emeritus of biochemistry, and director emeritus of the Beckman Center of Molecular and Genetic Medicine, at Stanford University, 279 Campus Drive, Stanford, California 94305, USA. He received the Nobel Prize in Chemistry in 1980 and the US National Medal of Science in 1983.
e-mail: pberg@stanford.edu

For more Meetings that Changed the World, see www.nature.com/nature/focus/meetings

NEWS & VIEWS

CANCER

Hedgehog's other great trick

Tom Curran and Jessica M. Y. Ng

The poet Archilochus wrote “The fox has many tricks, and the hedgehog one great one”. Yet a signalling-pathway namesake of hedgehog may have two great tricks that could be exploited for cancer therapy.

Molecular signalling pathways that control the growth of cancer cells have been heralded as prime targets for the treatment of a broad spectrum of tumours¹. Of particular interest is the hedgehog pathway, which normally is involved in development but whose abnormal activity is thought to drive the proliferation of cancer cells. In a few tumour types, this activity is known to occur by the mutation of genes encoding components of the hedgehog pathway, but it has also been proposed to underlie many other human cancers that do not contain mutations in the hedgehog-pathway genes. On page 406 of this issue, Yauch *et al.*² turn this model on its head by revealing the hedgehog pathway's other ‘clever trick’ in mediating cancer. They find that, in most cases, the oncogenic function of hedgehog proteins is not exerted in tumour cells. Instead, the proteins are secreted to activate the hedgehog pathway in non-malignant stromal cells, which, as part of the tumour microenvironment, support tumour growth.

In mice and humans, mutation of a gene called *Patched1* activates the hedgehog pathway by increasing the activity of smoothened (SMO), a cell-membrane protein required for hedgehog signalling. This leads to skin and brain tumours³. Inhibitors of SMO, including a small synthetic molecule called HhAntag, are highly effective in treating cancers such as basal-cell carcinoma and medulloblastoma in mouse disease models^{4,5}. The optimism generated by these animal studies was boosted by the results of a recent phase I clinical trial⁶, which reported that eight out of nine patients with advanced, metastatic basal-cell carcinoma showed a partial response to a synthetic SMO inhibitor, GDC-0449. The fact that cyclopamine (a natural SMO inhibitor) blocks even the growth of human tumour cell lines that lack mutations in the hedgehog pathway^{7,8} raised hopes for

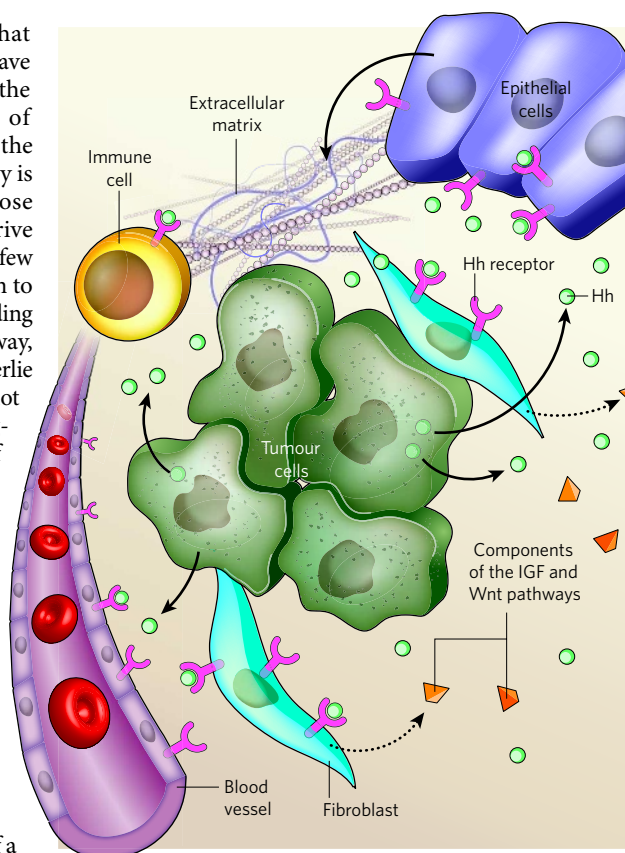


Figure 1 | Two-way communication between tumour cells and their microenvironment. Yauch *et al.*² find that many tumours secrete hedgehog-pathway activators (Hh) that act, through the SMO protein, on stromal cells in the tumour microenvironment (including blood-vessel cells, epithelial cells, fibroblasts and immune cells). This results in indirect support for tumour growth and survival through mechanisms originating in the stromal cells. Such mechanisms could include physical support of tumour cells, effects on the extracellular matrix, and secretion of components of molecular signalling pathways involving insulin-like growth factor (IGF) and Wnt.

this therapeutic strategy to giddy heights.

Yauch *et al.*² dispute this last conclusion, and deflate expectations to more realistic levels. Nonetheless, their findings open up the possibility that hedgehog inhibitors might be used to treat a broad range of tumours. Taking an industrial-scale approach, the authors tested hedgehog inhibitors in 125 human tumour

cell lines, but found no correlation between drug sensitivity and activity of the hedgehog pathway. They found instead that inhibition of tumour-cell growth by cyclopamine or HhAntag seems to be a consequence of off-target effects at high concentrations of these drugs.

In contrast to normal cells, none of the tumour lines responded to hedgehog proteins. This observation is consistent with a previous finding⁹ that culturing mouse medulloblastoma cells results in loss of hedgehog-pathway activity. But Yauch and colleagues also confirmed the presence of hedgehog proteins in several cancers. So the question remained: what were the hedgehog proteins doing there?

The answer came from studies of 25 human tumours transplanted directly into mice. This classic cancer model, called a xenograft, has been used extensively to investigate responses to anti-cancer drugs. It has a spotted history, however, because its predictions of a drug's clinical activity frequently don't come true. To avoid any problems associated with using cultured tumour cells for the xenografts, Yauch *et al.* used freshly generated xenografts.

When human tumours are transplanted into mice, stroma of mouse origin rapidly replaces human stromal cells¹⁰. To distinguish the responses of genes of the hedgehog pathway in human tumour cells from those in the mouse stroma, the authors measured gene-expression levels using species-specific markers. They found that, remarkably, levels of expression of genes encoding hedgehog-pathway activators in the tumours do not correlate with their expression levels in human tumour cells; instead, they correlate with gene expression levels in the mouse stroma. Moreover, treatment of mice with HhAntag results in a decreased stromal response — that is, reduced expression of hedgehog-pathway target genes — and a concomitant reduction in tumour growth².

To confirm the functional role of the hedgehog pathway in stroma, Yauch and colleagues used a genetic approach in which the gene encoding SMO was ablated from mouse embryonic fibroblast cells that were then co-injected with tumour cells to serve as a source of stromal support. Reduced SMO levels in stromal cells resulted in much smaller tumours. Finally, hedgehog stimulation of stroma increased the expression of genes involved in two other signalling pathways — those involving insulin-like growth factor and the Wnt protein — that can potentially also support cancer-cell growth.

So hedgehog inhibitors disrupt the intimate communication processes between tumour cells and their microenvironment. This seems reminiscent of the normal developmental role of hedgehog — mediating a dialogue between epithelial cells and their supporting mesenchymal cells. Yauch *et al.* argue that tumour cells thrive by usurping these normal developmental processes. Indeed, one can envisage a network of dependencies involving secreted or cell-surface proteins that is maintained by continuous stimulation of the hedgehog pathway (Fig. 1). Thus, attacking SMO strikes at a linchpin in tumour biology. This is an attractive concept, and one that points to alternative therapeutic targets. Targeting the stroma to treat cancer is not a new idea. For decades, Judah Folkman championed this concept in the form of anti-angiogenic therapies, which target tumour vasculature¹¹. Nonetheless, several uncertainties must be resolved before the hedgehog pathway can be routinely targeted for treating cancer.

First, it is not known exactly which stromal cell types and mechanisms support tumour growth (Fig. 1), and it could be that hedgehog influences several events in the tumour microenvironment. Yauch *et al.*² clearly show that HhAntag reduces the growth of various human tumours (with or without mutations in the hedgehog pathway) transplanted into mice. The effect they observe, however, is less dramatic than that seen^{4,5} when mouse tumours carrying intrinsic hedgehog-pathway mutations are eliminated. For most cancers, therefore, HhAntag may be more appropriate as an adjunct therapy. This would provide a powerful one-two punch, in which HhAntag targeting the stroma complements a primary therapy directed against tumour cells.

Another unresolved problem is how to identify the potentially large class of tumours that lack mutations in the hedgehog pathway and yet could benefit from such treatment. Expression of hedgehog proteins by tumours is a good indicator of susceptibility, but not an absolute criterion. Understanding the biology of tumour communication with its microenvironment would help in making such assessments. The good news is that both of hedgehog's great tricks (intrinsic and extrinsic signalling in tumours) can be exploited by HhAntag for the benefit of patients with cancer. ■

Tom Curran and Jessica M. Y. Ng are at the Joseph Stokes Jr. Research Institute, The Children's Hospital of Philadelphia, 517 Abramson Research Center, Philadelphia, Pennsylvania 19104-4318, USA.
e-mails: currant@chop.edu; ngj@chop.edu

1. Lobo, N. A., Shimon, Y., Qian, D. & Clarke, M. F. *Annu. Rev. Cell Dev. Biol.* **23**, 675–699 (2007).
2. Yauch, R. L. *et al.* *Nature* **455**, 406–410 (2008).
3. Rubin, L. L. & de Sauvage, F. J. *Nature Rev. Drug*

Discov. **5**, 1026–1033 (2006).

4. Williams, J. A. *et al.* *Proc. Natl Acad. Sci. USA* **100**, 4616–4621 (2003).
5. Romer, J. T. *et al.* *Cancer Cell* **6**, 229–240 (2004).
6. Von Hoff, D. D. *et al.* *Proc. 2008 Annu. Meet. Am. Assoc. Cancer Res. Abstr.* LB-138 (2008).
7. Karhadkar, S. S. *et al.* *Nature* **431**, 707–712 (2004).
8. Clement, V. *et al.* *Curr. Biol.* **17**, 165–172 (2007).
9. Sasai, K. *et al.* *Cancer Res.* **66**, 4215–4222 (2006).
10. Duda, D. G. *et al.* *Cancer Res.* **64**, 5920–5924 (2004).
11. Cooke, R. Dr. Folkman's War: *Angiogenesis and the Struggle to Defeat Cancer* (Random House, 2001).

ASTROPHYSICS

Heartbeat of an active galaxy

Phil Uttley

Active galaxies are thought to be the supermassive cousins of binary systems in our Galaxy that contain black holes. Observations of an unusual active galaxy provide compelling evidence that these systems are indeed related.

“Black holes have no hair.” This famous comment by John Wheeler encapsulates the idea that black holes are inherently simple objects, specified only by their mass, spin and electric charge. It also extends to their nearby environments. The superheated, X-ray-emitting gas that is pulled in, or ‘accreted’, from a black hole’s surroundings is predicted to show the same behaviour whether it is torn from a normal star by a ten-solar-mass black-hole companion — a system known as an X-ray binary — or accreted from the innermost regions of an ‘active’ galaxy by a central, supermassive black hole (of 10^6 – 10^9 solar masses). This prediction has been borne out by the discovery of many features common to X-ray binaries and their supermassive cousins, including powerful jets of radio-emitting plasma¹, relativistically broadened fluorescent emission lines² and characteristic variability timescales that scale with black-hole mass³.

But one feature has been missing. For more than 20 years, X-ray binaries have been known to show rapid, almost periodic oscillations in their X-ray emission⁴, but no such ‘quasi-periodic oscillations’ (QPOs) have been seen in the otherwise similar X-ray emissions of active galaxies. Observations by the XMM-Newton observatory, reported by Gierliński *et al.*⁵ on page 369 of this issue, now reveal the first X-ray QPO in an active galaxy, providing a further link between black holes across the mass scale.

Claims of QPOs in active galaxies have been made before, but this is the first one in which astronomers can have real confidence. The best previous claim was made from data obtained with a non-imaging detector, and was later found to be caused by X-rays emitted by a star that also exhibits QPOs (a cataclysmic variable star), and that happened to lie in the same field of view as the target active galaxy⁶. By contrast, XMM-Newton is an imaging instrument, and

the QPO reported by Gierliński *et al.*⁵ comes unambiguously from the target active galaxy, RE J1034+396. Other QPO claims were rejected⁷ after more careful tests that took into account the presence of flicker-noise-type fluctuations, which are ubiquitous in the X-ray emission from active galaxies. The QPO in RE J1034+396 passes these tests with flying colours, largely because the oscillation is highly coherent, showing almost exactly periodic variations over 16 of the 24 roughly 1-hour cycles of oscillation that could be observed. The detection is also helped by the apparent weakness of the flicker noise, with most of the variability being concentrated in the QPO and not over a broad range of timescales.

The discovery⁵ of the first QPO in an active galaxy extends the similarity of stellar-mass and supermassive black holes to new territory. But interpreting the QPO is difficult. A whole array of QPO periods, covering a wide range of timescales, is seen in X-ray binaries — which one, if any, corresponds to the QPO in RE J1034+396? To identify it, one would need to know the mass of the black hole, which acts as a scaling factor in determining the rate of oscillation. But this is not known for RE J1034+396.

Even more difficult is the question of the origin of QPOs. Despite many examples, and decades of study of QPOs in X-ray binaries, astronomers still don't know what causes them. Theorists have long suspected that they relate to characteristic timescales in the flow of plasma on to the black hole. A good candidate would be orbital timescales in the spinning accretion disk, which channels most of the plasma on to the black hole in a well-defined flow that may be capable of producing coherent oscillations. However, this picture suffers from a major drawback — QPOs in X-ray binaries seem to be strongest in the emission from the hottest plasma, which is probably not confined to the

To confirm the functional role of the hedgehog pathway in stroma, Yauch and colleagues used a genetic approach in which the gene encoding SMO was ablated from mouse embryonic fibroblast cells that were then co-injected with tumour cells to serve as a source of stromal support. Reduced SMO levels in stromal cells resulted in much smaller tumours. Finally, hedgehog stimulation of stroma increased the expression of genes involved in two other signalling pathways — those involving insulin-like growth factor and the Wnt protein — that can potentially also support cancer-cell growth.

So hedgehog inhibitors disrupt the intimate communication processes between tumour cells and their microenvironment. This seems reminiscent of the normal developmental role of hedgehog — mediating a dialogue between epithelial cells and their supporting mesenchymal cells. Yauch *et al.* argue that tumour cells thrive by usurping these normal developmental processes. Indeed, one can envisage a network of dependencies involving secreted or cell-surface proteins that is maintained by continuous stimulation of the hedgehog pathway (Fig. 1). Thus, attacking SMO strikes at a linchpin in tumour biology. This is an attractive concept, and one that points to alternative therapeutic targets. Targeting the stroma to treat cancer is not a new idea. For decades, Judah Folkman championed this concept in the form of anti-angiogenic therapies, which target tumour vasculature¹¹. Nonetheless, several uncertainties must be resolved before the hedgehog pathway can be routinely targeted for treating cancer.

First, it is not known exactly which stromal cell types and mechanisms support tumour growth (Fig. 1), and it could be that hedgehog influences several events in the tumour microenvironment. Yauch *et al.*² clearly show that HhAntag reduces the growth of various human tumours (with or without mutations in the hedgehog pathway) transplanted into mice. The effect they observe, however, is less dramatic than that seen^{4,5} when mouse tumours carrying intrinsic hedgehog-pathway mutations are eliminated. For most cancers, therefore, HhAntag may be more appropriate as an adjunct therapy. This would provide a powerful one-two punch, in which HhAntag targeting the stroma complements a primary therapy directed against tumour cells.

Another unresolved problem is how to identify the potentially large class of tumours that lack mutations in the hedgehog pathway and yet could benefit from such treatment. Expression of hedgehog proteins by tumours is a good indicator of susceptibility, but not an absolute criterion. Understanding the biology of tumour communication with its microenvironment would help in making such assessments. The good news is that both of hedgehog's great tricks (intrinsic and extrinsic signalling in tumours) can be exploited by HhAntag for the benefit of patients with cancer. ■

Tom Curran and Jessica M. Y. Ng are at the Joseph Stokes Jr. Research Institute, The Children's Hospital of Philadelphia, 517 Abramson Research Center, Philadelphia, Pennsylvania 19104-4318, USA.
e-mails: currant@chop.edu; ngj@chop.edu

1. Lobo, N. A., Shimono, Y., Qian, D. & Clarke, M. F. *Annu. Rev. Cell Dev. Biol.* **23**, 675–699 (2007).
2. Yauch, R. L. *et al.* *Nature* **455**, 406–410 (2008).
3. Rubin, L. L. & de Sauvage, F. J. *Nature Rev. Drug*

Discov. **5**, 1026–1033 (2006).

4. Williams, J. A. *et al.* *Proc. Natl Acad. Sci. USA* **100**, 4616–4621 (2003).
5. Romer, J. T. *et al.* *Cancer Cell* **6**, 229–240 (2004).
6. Von Hoff, D. D. *et al.* *Proc. 2008 Annu. Meet. Am. Assoc. Cancer Res. Abstr.* LB-138 (2008).
7. Karhadkar, S. S. *et al.* *Nature* **431**, 707–712 (2004).
8. Clement, V. *et al.* *Curr. Biol.* **17**, 165–172 (2007).
9. Sasai, K. *et al.* *Cancer Res.* **66**, 4215–4222 (2006).
10. Duda, D. G. *et al.* *Cancer Res.* **64**, 5920–5924 (2004).
11. Cooke, R. Dr. Folkman's War: *Angiogenesis and the Struggle to Defeat Cancer* (Random House, 2001).

ASTROPHYSICS

Heartbeat of an active galaxy

Phil Uttley

Active galaxies are thought to be the supermassive cousins of binary systems in our Galaxy that contain black holes. Observations of an unusual active galaxy provide compelling evidence that these systems are indeed related.

“Black holes have no hair.” This famous comment by John Wheeler encapsulates the idea that black holes are inherently simple objects, specified only by their mass, spin and electric charge. It also extends to their nearby environments. The superheated, X-ray-emitting gas that is pulled in, or ‘accreted’, from a black hole’s surroundings is predicted to show the same behaviour whether it is torn from a normal star by a ten-solar-mass black-hole companion — a system known as an X-ray binary — or accreted from the innermost regions of an ‘active’ galaxy by a central, supermassive black hole (of 10^6 – 10^9 solar masses). This prediction has been borne out by the discovery of many features common to X-ray binaries and their supermassive cousins, including powerful jets of radio-emitting plasma¹, relativistically broadened fluorescent emission lines² and characteristic variability timescales that scale with black-hole mass³.

But one feature has been missing. For more than 20 years, X-ray binaries have been known to show rapid, almost periodic oscillations in their X-ray emission⁴, but no such ‘quasi-periodic oscillations’ (QPOs) have been seen in the otherwise similar X-ray emissions of active galaxies. Observations by the XMM-Newton observatory, reported by Gierliński *et al.*⁵ on page 369 of this issue, now reveal the first X-ray QPO in an active galaxy, providing a further link between black holes across the mass scale.

Claims of QPOs in active galaxies have been made before, but this is the first one in which astronomers can have real confidence. The best previous claim was made from data obtained with a non-imaging detector, and was later found to be caused by X-rays emitted by a star that also exhibits QPOs (a cataclysmic variable star), and that happened to lie in the same field of view as the target active galaxy⁶. By contrast, XMM-Newton is an imaging instrument, and

the QPO reported by Gierliński *et al.*⁵ comes unambiguously from the target active galaxy, RE J1034+396. Other QPO claims were rejected⁷ after more careful tests that took into account the presence of flicker-noise-type fluctuations, which are ubiquitous in the X-ray emission from active galaxies. The QPO in RE J1034+396 passes these tests with flying colours, largely because the oscillation is highly coherent, showing almost exactly periodic variations over 16 of the 24 roughly 1-hour cycles of oscillation that could be observed. The detection is also helped by the apparent weakness of the flicker noise, with most of the variability being concentrated in the QPO and not over a broad range of timescales.

The discovery⁵ of the first QPO in an active galaxy extends the similarity of stellar-mass and supermassive black holes to new territory. But interpreting the QPO is difficult. A whole array of QPO periods, covering a wide range of timescales, is seen in X-ray binaries — which one, if any, corresponds to the QPO in RE J1034+396? To identify it, one would need to know the mass of the black hole, which acts as a scaling factor in determining the rate of oscillation. But this is not known for RE J1034+396.

Even more difficult is the question of the origin of QPOs. Despite many examples, and decades of study of QPOs in X-ray binaries, astronomers still don't know what causes them. Theorists have long suspected that they relate to characteristic timescales in the flow of plasma on to the black hole. A good candidate would be orbital timescales in the spinning accretion disk, which channels most of the plasma on to the black hole in a well-defined flow that may be capable of producing coherent oscillations. However, this picture suffers from a major drawback — QPOs in X-ray binaries seem to be strongest in the emission from the hottest plasma, which is probably not confined to the

disk. The QPO in RE J1034+396 also seems to follow this trend.

More detailed simulations of the gas flow and interactions between the disk and hot plasma, including radiation effects, may help to solve these problems. But such advances are some way off. Even so, the study of QPOs in active galaxies such as RE J1034+396 could provide a unique insight into their origin. This is because the QPOs from supermassive black holes are much slower than their counterparts in X-ray binaries, so that many more X-ray photons can be accumulated over each cycle of the oscillation. This means that the QPOs in active galaxies can be studied in exquisite slow-motion detail, which should surely constrain models for their origin and shed light on conditions close to the black hole, where the oscillations are likely to be produced. The QPO in RE J1034+396 already suggests that these fluctuations can be much more coherent over short timescales than was previously thought possible on the basis of observations of X-ray-binary QPOs.

Although more detailed X-ray observations of the QPO in RE J1034+396 are warranted, further examples of QPOs in other active

galaxies will be required to make broader progress. They may, however, be rare. Gierliński *et al.*⁵ note that RE J1034+396 is an unusual active galaxy, with a broadband spectrum that suggests an extremely high rate of accretion of gas onto the black hole. It seems unlikely to be a coincidence that the first active-galaxy QPO has been found in one of the most extreme types of these galaxies. Perhaps there is something about the exceptional fuelling rate in RE J1034+396 that favours QPO production. If so, that might make further examples difficult to find. ■

Phil Uttley is at the School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
e-mail: pu@phys.soton.ac.uk

1. Mirabel, I. F. & Rodríguez, L. F. *Nature* **371**, 46–48 (1994).
2. Miller, J. M. *et al.* *Astrophys. J.* **606**, L131–L134 (2004).
3. M'Hardy, I. M., Koording, E., Knigge, C., Uttley, P. & Fender, R. P. *Nature* **444**, 730–732 (2006).
4. van der Klis, M. *et al.* *Nature* **316**, 225–230 (1985).
5. Gierliński, M., Middleton, M., Ward, M. & Done, C. *Nature* **455**, 369–371 (2008).
6. Madejski, G. M. *et al.* *Nature* **365**, 626–628 (1993).
7. Vaughan, S. & Uttley, P. *Adv. Space Res.* **38**, 1405–1408 (2006).

DEVELOPMENTAL NEUROSCIENCE

Hox and Fox

Samuel L. Pfaff

A network of Hox factors specifies the identity of motor neurons in the spinal cord. Studies of an essential Fox-class accessory factor illuminate the developmental and evolutionary subtlety of the process.

The vertebrate central nervous system (CNS) contains many classes of neurons that have specialized properties. Imagine the problem of creating this cellular diversity, in which neurons that look alike have in fact acquired fine-grained distinctions to ensure that the proper synaptic connections with other neurons are established and maintained. There are an estimated 100 trillion synapses in the brains of the most complex vertebrates, and neurobiologists are finding that nature has devised a multitude of clever strategies to guide neuron development and connectivity.

Remarkable inroads into understanding these strategies have been made by examining the development of a relatively simple part of the CNS — the spinal cord¹. Most notably, it has been discovered that the evolutionarily conserved Hox family of gene transcription factors operates in a sophisticated cross-regulatory network to specify individual subtypes of motor neurons that are preordained to connect with specific muscles and ganglia in the periphery of an animal². Work published in *Cell* (Dasen *et al.*)³ and *Neuron* (Roussou *et al.*)⁴ peels away a further layer of the developmental onion. The two papers report that another

transcription factor, called FoxP1, helps Hox proteins to regulate the genes that control motor-neuron diversification.

Motor neurons represent the final command stage in a chain of cells extending from the brain to the spinal cord. They trigger activity in the autonomic (sympathetic) nervous system, such as sweating under stressful conditions, and they send instructions to muscles in the form of the neurotransmitter molecule acetylcholine that cause the muscles to contract. Motor coordination is achieved by the precise spatial and temporal control of hundreds of muscles. Consequently, the complexity of the musculoskeletal system is mirrored by a similarly complex arrangement of motor neurons in the spinal cord.

During embryo formation, motor neurons grow axons to their appropriate muscles in the periphery — the equivalent in IT terms of threading Internet cable to its destination. Among the motor-neuron subclasses examined by Dasen *et al.*³ and Roussou *et al.*⁴ were the median motor column (MMC) neurons that innervate axial muscles supporting posture; the hypaxial motor column (HMC) neurons that innervate body-wall muscles around

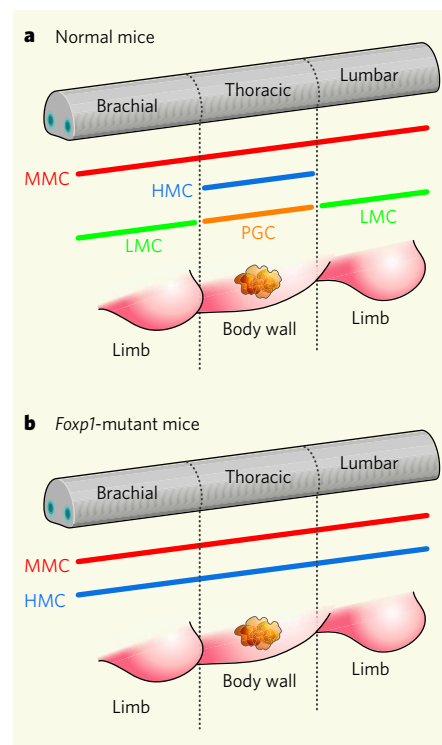


Figure 1 | Neuron specification in the spinal cord of embryonic mice^{3,4}. **a**, In normal mice, Hox genes and *Foxp1* interact to generate the correct subclass of motor neuron at the brachial, thoracic and lumbar levels of the spinal cord. The cell bodies of the neurons align into longitudinal columns called the LMC, which innervates limbs; MMC, which innervates axial muscles; HMC, which innervates body-wall muscles; and PGC, which innervates ganglia of the sympathetic nervous system. **b**, Mutations in *Foxp1* disrupt Hox activity involved in specifying the LMC and PGC. The spinal cord of *Foxp1* mutant mice defaults to generating motor neurons such as those of more primitive vertebrates, for example the lamprey. (Based on Fig. 1 of ref. 3.)

the rib cage; the preganglionic motor column (PGC) neurons that innervate peripheral ganglia of the autonomic nervous system; and the lateral motor column (LMC) neurons that innervate muscles in the arms and legs. The motor neurons in the spinal cord and the muscles they control are typically in register with one another (Fig. 1a). LMC neurons occur only in the brachial and lumbar spinal cord, aligned with the arms and legs respectively (or wings/legs in birds and fins and fish), whereas HMC neurons are located at thoracic levels where the rib cage is present.

The spinal cord is derived from undifferentiated cells that comprise the embryo's neural tube. The generation of specific subtypes of motor neuron is initiated by cell–cell interactions mediated by different secreted factors that in effect create an *x*, *y*, *z* coordinate system^{1–3}. In general, these signals activate intracellular pathways that eventually culminate in the transcription or repression of genes needed for the specialized properties of each type of motor neuron. Multiple layers

disk. The QPO in RE J1034+396 also seems to follow this trend.

More detailed simulations of the gas flow and interactions between the disk and hot plasma, including radiation effects, may help to solve these problems. But such advances are some way off. Even so, the study of QPOs in active galaxies such as RE J1034+396 could provide a unique insight into their origin. This is because the QPOs from supermassive black holes are much slower than their counterparts in X-ray binaries, so that many more X-ray photons can be accumulated over each cycle of the oscillation. This means that the QPOs in active galaxies can be studied in exquisite slow-motion detail, which should surely constrain models for their origin and shed light on conditions close to the black hole, where the oscillations are likely to be produced. The QPO in RE J1034+396 already suggests that these fluctuations can be much more coherent over short timescales than was previously thought possible on the basis of observations of X-ray-binary QPOs.

Although more detailed X-ray observations of the QPO in RE J1034+396 are warranted, further examples of QPOs in other active

galaxies will be required to make broader progress. They may, however, be rare. Gierliński *et al.*⁵ note that RE J1034+396 is an unusual active galaxy, with a broadband spectrum that suggests an extremely high rate of accretion of gas onto the black hole. It seems unlikely to be a coincidence that the first active-galaxy QPO has been found in one of the most extreme types of these galaxies. Perhaps there is something about the exceptional fuelling rate in RE J1034+396 that favours QPO production. If so, that might make further examples difficult to find. ■

Phil Uttley is at the School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK.
e-mail: pu@phys.soton.ac.uk

1. Mirabel, I. F. & Rodríguez, L. F. *Nature* **371**, 46–48 (1994).
2. Miller, J. M. *et al.* *Astrophys. J.* **606**, L131–L134 (2004).
3. M'Hardy, I. M., Koording, E., Knigge, C., Uttley, P. & Fender, R. P. *Nature* **444**, 730–732 (2006).
4. van der Klis, M. *et al.* *Nature* **316**, 225–230 (1985).
5. Gierliński, M., Middleton, M., Ward, M. & Done, C. *Nature* **455**, 369–371 (2008).
6. Madejski, G. M. *et al.* *Nature* **365**, 626–628 (1993).
7. Vaughan, S. & Uttley, P. *Adv. Space Res.* **38**, 1405–1408 (2006).

DEVELOPMENTAL NEUROSCIENCE

Hox and Fox

Samuel L. Pfaff

A network of Hox factors specifies the identity of motor neurons in the spinal cord. Studies of an essential Fox-class accessory factor illuminate the developmental and evolutionary subtlety of the process.

The vertebrate central nervous system (CNS) contains many classes of neurons that have specialized properties. Imagine the problem of creating this cellular diversity, in which neurons that look alike have in fact acquired fine-grained distinctions to ensure that the proper synaptic connections with other neurons are established and maintained. There are an estimated 100 trillion synapses in the brains of the most complex vertebrates, and neurobiologists are finding that nature has devised a multitude of clever strategies to guide neuron development and connectivity.

Remarkable inroads into understanding these strategies have been made by examining the development of a relatively simple part of the CNS — the spinal cord¹. Most notably, it has been discovered that the evolutionarily conserved Hox family of gene transcription factors operates in a sophisticated cross-regulatory network to specify individual subtypes of motor neurons that are preordained to connect with specific muscles and ganglia in the periphery of an animal². Work published in *Cell* (Dasen *et al.*)³ and *Neuron* (Roussou *et al.*)⁴ peels away a further layer of the developmental onion. The two papers report that another

transcription factor, called FoxP1, helps Hox proteins to regulate the genes that control motor-neuron diversification.

Motor neurons represent the final command stage in a chain of cells extending from the brain to the spinal cord. They trigger activity in the autonomic (sympathetic) nervous system, such as sweating under stressful conditions, and they send instructions to muscles in the form of the neurotransmitter molecule acetylcholine that cause the muscles to contract. Motor coordination is achieved by the precise spatial and temporal control of hundreds of muscles. Consequently, the complexity of the musculoskeletal system is mirrored by a similarly complex arrangement of motor neurons in the spinal cord.

During embryo formation, motor neurons grow axons to their appropriate muscles in the periphery — the equivalent in IT terms of threading Internet cable to its destination. Among the motor-neuron subclasses examined by Dasen *et al.*³ and Roussou *et al.*⁴ were the median motor column (MMC) neurons that innervate axial muscles supporting posture; the hypaxial motor column (HMC) neurons that innervate body-wall muscles around

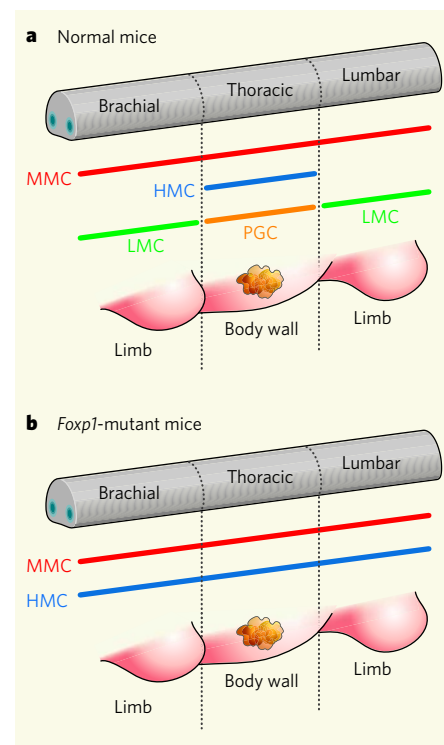


Figure 1 | Neuron specification in the spinal cord of embryonic mice^{3,4}. **a**, In normal mice, Hox genes and *Foxp1* interact to generate the correct subclass of motor neuron at the brachial, thoracic and lumbar levels of the spinal cord. The cell bodies of the neurons align into longitudinal columns called the LMC, which innervates limbs; MMC, which innervates axial muscles; HMC, which innervates body-wall muscles; and PGC, which innervates ganglia of the sympathetic nervous system. **b**, Mutations in *Foxp1* disrupt Hox activity involved in specifying the LMC and PGC. The spinal cord of *Foxp1* mutant mice defaults to generating motor neurons such as those of more primitive vertebrates, for example the lamprey. (Based on Fig. 1 of ref. 3.)

the rib cage; the preganglionic motor column (PGC) neurons that innervate peripheral ganglia of the autonomic nervous system; and the lateral motor column (LMC) neurons that innervate muscles in the arms and legs. The motor neurons in the spinal cord and the muscles they control are typically in register with one another (Fig. 1a). LMC neurons occur only in the brachial and lumbar spinal cord, aligned with the arms and legs respectively (or wings/legs in birds and fins and fish), whereas HMC neurons are located at thoracic levels where the rib cage is present.

The spinal cord is derived from undifferentiated cells that comprise the embryo's neural tube. The generation of specific subtypes of motor neuron is initiated by cell–cell interactions mediated by different secreted factors that in effect create an *x*, *y*, *z* coordinate system^{1–3}. In general, these signals activate intracellular pathways that eventually culminate in the transcription or repression of genes needed for the specialized properties of each type of motor neuron. Multiple layers

of transcription factors are then involved in establishing motor-neuron identity^{1,5}, with the Hox family lying at the core of the circuitry. Hox genes have undergone several duplications in more highly evolved vertebrates, presumably to expand their repertoire of functions. In mice, for example, Hox6 proteins contribute to the specification of brachial LMC motor neurons, Hox9 specify PGC neurons, and Hox10 specify lumbar LMC neurons.

In some cases, however, Hox expression imprecisely correlates with motor-neuron subtype. Dasen *et al.*³ and Roussou *et al.*⁴ set out to understand why, starting with the knowledge that Hox proteins typically function with cofactors such as Meis, Pbx/Prep, Engrailed and Fox^{6,7}. Meis and Pbx/Prep are not restricted to particular subtypes, and Engrailed is not expressed at all by motor neurons, so these were poor candidates for further study. In contrast, FoxP1 looked promising, being expressed at low levels in PGC motor neurons and at high levels in LMC cells.

Both groups^{3,4} found that mutation of the mouse *Foxp1* gene does not interfere with motor-neuron generation as such, but that PGC and LMC cells fail to emerge and motor neurons seem instead to remain stuck in their 'ground state' and resemble HMC cells. From studies with the fruitfly *Drosophila*, we can infer that Hox proteins functionally interact with the Fox homologue called sloppy paired (Slp)^{6,7}, which could modify their DNA-binding and/or transcriptional activity. Nevertheless, Hox proteins also seem to have FoxP1-independent functions³, and therefore individual genes in each motor-neuron class could contain either Hox-only or Hox-FoxP1 complexes, which as a composite might specify subtype identity.

Furthermore, the concentration of FoxP1 seems to be significant: misexpression of high levels favours LMC development, whereas low levels trigger PGC specification³. The mechanistic basis for this observation remains unknown, but one possibility is that ternary Hox complexes form with different amounts of FoxP1 and have unique activities. Understanding these details might help us to understand how FoxP1 also functions in non-neuronal contexts such as cardiac- and blood-cell development.

Mice lacking *Foxp1* are not viable^{3,4} and have a strange type of spinal cord that in some ways resembles the spinal cord of more primitive animals (Fig. 1b). Ancient vertebrates such as lamprey and hagfish swim using axial and hypaxial muscles, and their spinal cords seem to have neurons of MMC and HMC character but not PGC and LMC cells⁸. The appearance of LMC and PGC neurons is linked to the formation of lateral fins and limbs and a sympathetic nervous system — features that emerged later in vertebrate evolution. Although the Hox and Fox genes can be traced back to invertebrates, evolutionary pressure evidently exploited them to expand the repertoire of functions mediated by the CNS. Comparisons of the promoters of

Fox genes and of Fox protein structures from different animals should help us to understand how genes can be co-opted from one context to be used in another⁹.

Samuel L. Pfaff is at the Howard Hughes Medical Institute, Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. e-mail: pfaff@salk.edu

1. Jessell, T. M. *Nature Rev. Genet.* **1**, 20–29 (2000).

2. Dasen, J. S., Tice, B. C., Brenner-Morton, S. & Jessell, T. M. *Cell* **123**, 477–491 (2005).
3. Dasen, J. S., De Camilli, A., Wang, B., Tucker, P. W. & Jessell, T. M. *Cell* **134**, 304–316 (2008).
4. Roussou, D. L., Gaber, Z. B., Wellik, D., Morrissey, E. E. & Novitsch, B. G. *Neuron* **59**, 226–240 (2008).
5. Shirasaki, R. & Pfaff, S. L. *Annu. Rev. Neurosci.* **25**, 251–281 (2002).
6. Mann, R. S. & Affolter, M. *Curr. Opin. Genet. Dev.* **8**, 423–429 (1998).
7. Moens, C. B. & Selleri, L. *Dev. Biol.* **291**, 193–206 (2006).
8. Fetcho, J. R. *Brain Behav. Evol.* **40**, 82–97 (1992).
9. True, J. R. & Carroll, S. B. *Annu. Rev. Cell Dev. Biol.* **18**, 53–80 (2002).

PHOTONICS

Illumination for atomic movies

Claudio Pellegrini

Free-electron lasers could produce X-rays intense enough to make atomic-resolution movies. Initial designs are kilometres long, but a prototype working in the ultraviolet points a way to shorter machines.

Scientists have long dreamed of having a video camera with which to record individual atoms and their motions, and perhaps even to follow electrons during physical, chemical and biological processes. Radiation of extremely small wavelengths is needed to resolve atomic and sub-atomic details, and so attention has focused on the use of high-energy electrons to produce coherent beams of X-rays. This can be done using new types of lasers called free-electron lasers (FELs). However, current designs for these devices are enormous, with sizes on the scale of kilometres. Writing in *Nature Photonics*, Shintake *et al.*¹ report the results of the successful operation of a FEL that produces

ultraviolet (UV) light, and whose design could lead to more manageably sized instruments.

An X-ray free-electron laser (X-FEL) would allow scientists to view the molecular structure of a single protein or virus particle, and take atomic-scale motion pictures of chemical processes on the timescale of a few femtoseconds (the time it takes an electron to move around an atom). The world's first X-FEL, called the Linear Coherent Light Source (LCLS)², is already being built at the Stanford Linear Accelerator Center (SLAC) and will be completed in 2009. A similar instrument, called XFEL³, will be built at the DESY synchrotron facility in Hamburg in the next few years by a European

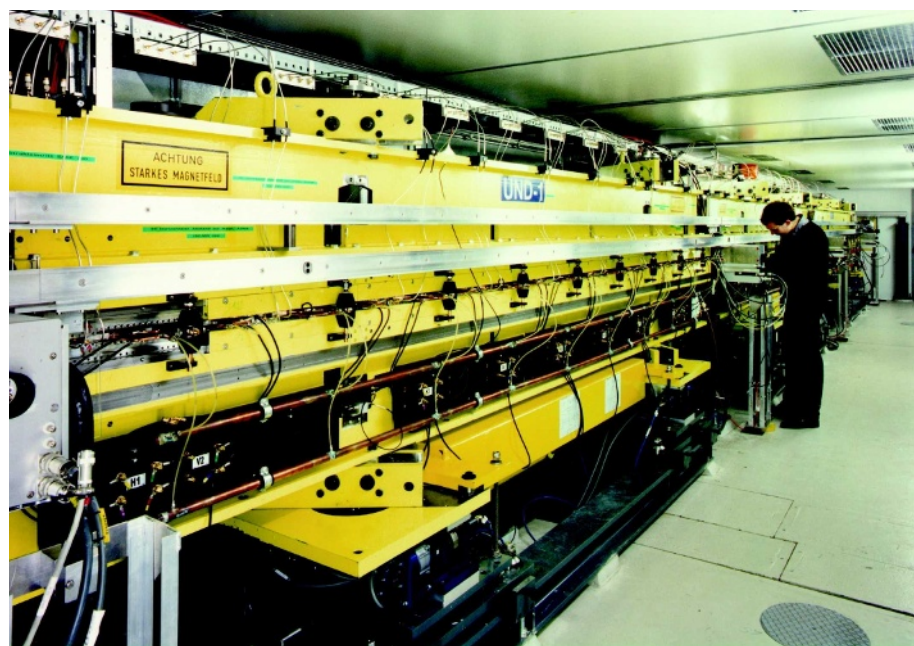


Figure 1 | Long lasers. Part of a free-electron laser (FEL) at the DESY synchrotron facility in Hamburg, Germany. This device produces light in the ultraviolet–soft-X-ray region of the electromagnetic spectrum, and is a prototype for an X-ray FEL under construction at DESY.

DESY

of transcription factors are then involved in establishing motor-neuron identity^{1,5}, with the Hox family lying at the core of the circuitry. Hox genes have undergone several duplications in more highly evolved vertebrates, presumably to expand their repertoire of functions. In mice, for example, Hox6 proteins contribute to the specification of brachial LMC motor neurons, Hox9 specify PGC neurons, and Hox10 specify lumbar LMC neurons.

In some cases, however, Hox expression imprecisely correlates with motor-neuron subtype. Dasen *et al.*³ and Roussou *et al.*⁴ set out to understand why, starting with the knowledge that Hox proteins typically function with cofactors such as Meis, Pbx/Prep, Engrailed and Fox^{6,7}. Meis and Pbx/Prep are not restricted to particular subtypes, and Engrailed is not expressed at all by motor neurons, so these were poor candidates for further study. In contrast, FoxP1 looked promising, being expressed at low levels in PGC motor neurons and at high levels in LMC cells.

Both groups^{3,4} found that mutation of the mouse *Foxp1* gene does not interfere with motor-neuron generation as such, but that PGC and LMC cells fail to emerge and motor neurons seem instead to remain stuck in their 'ground state' and resemble HMC cells. From studies with the fruitfly *Drosophila*, we can infer that Hox proteins functionally interact with the Fox homologue called sloppy paired (Slp)^{6,7}, which could modify their DNA-binding and/or transcriptional activity. Nevertheless, Hox proteins also seem to have FoxP1-independent functions³, and therefore individual genes in each motor-neuron class could contain either Hox-only or Hox-FoxP1 complexes, which as a composite might specify subtype identity.

Furthermore, the concentration of FoxP1 seems to be significant: misexpression of high levels favours LMC development, whereas low levels trigger PGC specification³. The mechanistic basis for this observation remains unknown, but one possibility is that ternary Hox complexes form with different amounts of FoxP1 and have unique activities. Understanding these details might help us to understand how FoxP1 also functions in non-neuronal contexts such as cardiac- and blood-cell development.

Mice lacking *Foxp1* are not viable^{3,4} and have a strange type of spinal cord that in some ways resembles the spinal cord of more primitive animals (Fig. 1b). Ancient vertebrates such as lamprey and hagfish swim using axial and hypaxial muscles, and their spinal cords seem to have neurons of MMC and HMC character but not PGC and LMC cells⁸. The appearance of LMC and PGC neurons is linked to the formation of lateral fins and limbs and a sympathetic nervous system — features that emerged later in vertebrate evolution. Although the Hox and Fox genes can be traced back to invertebrates, evolutionary pressure evidently exploited them to expand the repertoire of functions mediated by the CNS. Comparisons of the promoters of

Fox genes and of Fox protein structures from different animals should help us to understand how genes can be co-opted from one context to be used in another⁹.

Samuel L. Pfaff is at the Howard Hughes Medical Institute, Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, USA. e-mail: pfaff@salk.edu

1. Jessell, T. M. *Nature Rev. Genet.* **1**, 20–29 (2000).

2. Dasen, J. S., Tice, B. C., Brenner-Morton, S. & Jessell, T. M. *Cell* **123**, 477–491 (2005).
3. Dasen, J. S., De Camilli, A., Wang, B., Tucker, P. W. & Jessell, T. M. *Cell* **134**, 304–316 (2008).
4. Roussou, D. L., Gaber, Z. B., Wellik, D., Morrissey, E. E. & Novitsch, B. G. *Neuron* **59**, 226–240 (2008).
5. Shirasaki, R. & Pfaff, S. L. *Annu. Rev. Neurosci.* **25**, 251–281 (2002).
6. Mann, R. S. & Affolter, M. *Curr. Opin. Genet. Dev.* **8**, 423–429 (1998).
7. Moens, C. B. & Selleri, L. *Dev. Biol.* **291**, 193–206 (2006).
8. Fetcho, J. R. *Brain Behav. Evol.* **40**, 82–97 (1992).
9. True, J. R. & Carroll, S. B. *Annu. Rev. Cell Dev. Biol.* **18**, 53–80 (2002).

PHOTONICS

Illumination for atomic movies

Claudio Pellegrini

Free-electron lasers could produce X-rays intense enough to make atomic-resolution movies. Initial designs are kilometres long, but a prototype working in the ultraviolet points a way to shorter machines.

Scientists have long dreamed of having a video camera with which to record individual atoms and their motions, and perhaps even to follow electrons during physical, chemical and biological processes. Radiation of extremely small wavelengths is needed to resolve atomic and sub-atomic details, and so attention has focused on the use of high-energy electrons to produce coherent beams of X-rays. This can be done using new types of lasers called free-electron lasers (FELs). However, current designs for these devices are enormous, with sizes on the scale of kilometres. Writing in *Nature Photonics*, Shintake *et al.*¹ report the results of the successful operation of a FEL that produces

ultraviolet (UV) light, and whose design could lead to more manageably sized instruments.

An X-ray free-electron laser (X-FEL) would allow scientists to view the molecular structure of a single protein or virus particle, and take atomic-scale motion pictures of chemical processes on the timescale of a few femtoseconds (the time it takes an electron to move around an atom). The world's first X-FEL, called the Linear Coherent Light Source (LCLS)², is already being built at the Stanford Linear Accelerator Center (SLAC) and will be completed in 2009. A similar instrument, called XFEL³, will be built at the DESY synchrotron facility in Hamburg in the next few years by a European

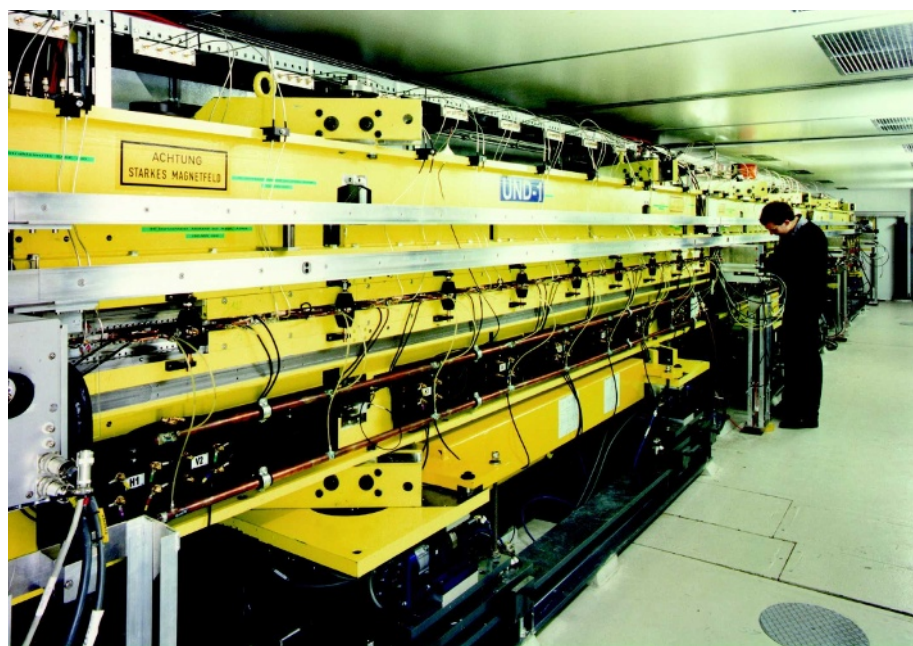


Figure 1 | Long lasers. Part of a free-electron laser (FEL) at the DESY synchrotron facility in Hamburg, Germany. This device produces light in the ultraviolet–soft-X-ray region of the electromagnetic spectrum, and is a prototype for an X-ray FEL under construction at DESY.

CLIMATE SCIENCE

A bright side of precipitation

During the period 1986–2000, the world — or at least its terrestrial parts — became wetter. On average, the increase in precipitation over this time was 3.5 mm yr^{-1} . In other words, the hydrological cycle intensified, and Martin Wild and his colleagues set themselves the task of investigating the factors responsible (M. Wild, J. Grieser and C. Schär *Geophys. Res. Lett.* **35**, L17706, doi:10.1029/2008GL034842; 2008).

The strength of the main engine that powers the hydrological cycle depends on the balance of radiative energy at Earth's surface. That balance in turn depends, first, on the amount of solar radiation absorbed by the surface, and second, on the difference between the amount of thermal radiation being transmitted

upwards into the atmosphere (upward radiation) and that being re-radiated back (downward radiation, which is controlled by the greenhouse capacity of the atmosphere). The sum of these fluxes determines the amount of energy that is available for the latent heat of evaporation, which drives precipitation. Any change in these fluxes induces changes in evaporation and precipitation, and thus in the intensity of the hydrological cycle.

For their calculations, the authors drew from two sources: the Global Energy Balance Archive and the Baseline Surface Radiation Network. The latter provides more sophisticated data, but started up only in the 1990s. Another consideration in the calculations was the serious but temporary



S. CURRY/AFP/GETTY IMAGES

effects on radiation balance of the huge emissions from the 1991 eruption of Mount Pinatubo in the Philippines, which lasted for several years.

After accounting for those effects, Wild *et al.* conclude that there has been a steady rise in net surface radiation of $0.2 \text{ W m}^{-2} \text{ yr}^{-1}$ between 1986 and 2000. They mention two possible causes. One is 'solar brightening' — a rise in surface radiation resulting from a more transparent atmosphere caused by a decrease in anthropogenic aerosols. The other is a more

powerful greenhouse effect, increasing downward radiation.

As the authors point out, their study cannot provide a truly global picture because of the lack of data from the oceans. But their calculated rise in net surface radiation fits well with the estimate of energy flux required to drive the rise in precipitation. This provides a satisfying first-order numerical link between the changes in precipitation and radiation, recorded in independent data sets, over the 15-year period of their study.

Tim Lincoln

collaboration (Fig. 1). For all their promise, the LCLS and XFEL are large and expensive; the LCLS is about 2 kilometres long, and XFEL will be even longer. The construction cost for the LCLS, which uses an existing particle accelerator, is about €300 million (\$450 million), and XFEL will cost about €1 billion.

Shintake and colleagues are part of a team designing a Japanese X-FEL called the SPring-8 Compact Self-Amplified Spontaneous Emission Source (SCSS), which is planned for construction in the near future⁴. The work they now describe¹ uses a smaller-scale FEL, working at UV wavelengths, to show that cheaper and smaller X-FELs can be made by taking advantage of advances in the physics and technology of electron beams.

FELs operate by combining the technology of lasers and electron accelerators (Fig. 2). A particle accelerator produces electrons moving at close to the speed of light. These pass through a magnet, which causes them to follow

a sinusoidal trajectory. The acceleration due to the bending of the electrons by this 'undulator' magnet produces electromagnetic waves. These have a wavelength that is proportional to the oscillation period of the sinusoidal path and inversely proportional to the square of the electron energy. Typically, the oscillation has a period of a few centimetres, and therefore, to produce X-rays with a wavelength of about 10^{-10} metres (1 ångström, the typical size of an atom), electrons are required that have energies of up to 15 gigaelectronvolts. For comparison, the energy of electrons produced by medical electron linear accelerators — used in many hospitals for radiation therapy — is 1,000 times smaller than this value. Medical accelerators are about a metre long; if an X-FEL accelerator were to be constructed using the same technology, it would need to be around a kilometre long. The length of the undulator magnet alone would be about 100 metres, and so, including space for X-ray diagnostic and manipulation

equipment, the entire X-FEL would measure a few kilometres.

One way to make X-FEL sources more manageable would be to reduce the electron-beam energy but to force the electrons into tighter undulations. This would have the effect of making both the accelerator and undulator shorter. Shintake *et al.*¹ show that the accelerator length can be reduced by using a 'C-band' high-frequency accelerator, a technology developed for the construction of high-energy electron–positron linear colliders⁵. Other advances involved are an improved electron source and an undulator magnet with smaller period and shorter length. When combined in the SCSS design, these advances would lead to a more compact X-FEL.

An even more advanced instrument under consideration^{6,7} is a high-gradient accelerator driven by a phenomenon known as a laser-plasma wakefield, which is created by passing a brief laser pulse through a plasma of charged particles. Whereas Shintake and colleagues' approach can decrease the length of the accelerator to around a half or a third of that possible now, the laser-plasma accelerator could provide a 100-fold reduction. However, it is still at the exploration and proof-of-principle stage.

The size and cost of an X-FEL could also be reduced by using sources that produce electron beams with higher electron density and smaller variation in momentum than are currently available; together, these properties are described as the beam emittance. Smaller-emittance electron beams would allow the use of lower electron energies, and shortened accelerators and undulators. Emittance could

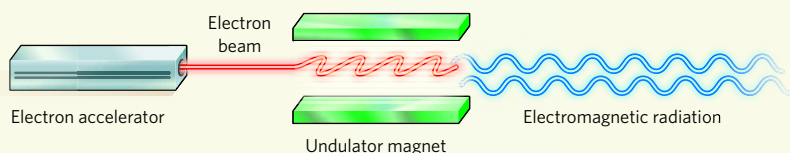


Figure 2 | X-ray specs. A free-electron laser (FEL) consists of a linear accelerator producing a beam of high-energy electrons that then pass through an 'undulator' magnet. This magnet bends the electrons into a sinusoidal path, causing them to emit electromagnetic radiation whose frequency depends on the energy of the electrons and the wavelength of their undulations. Using current technology, a FEL that emits X-rays would be several kilometres long, but the approach tested by Shintake *et al.*¹ on an ultraviolet-emitting device could reduce the size two- to threefold.

be reduced by optimizing the charge and volume of the bunches of electrons produced from the source⁸, a prediction confirmed by initial results at SLAC (P. Emma and D. Dowell, personal communication). Combining this approach with a high-gradient accelerator such as that used by Shintake *et al.*¹ would reduce the size and cost of FELs still further.

Because any reduction in cost and size also implies a reduction in the average number of photons produced, some experiments will always require large facilities such as the LCLS or XFEL. The new smaller systems would nevertheless add significantly to our experimental capabilities. There is a new world to explore at atomic and molecular scales of time and distance, and a multiplicity of approaches, such as that described by Shintake *et al.*, together with other FEL

systems, would speed up the exploration of this uncharted territory. ■

Claudio Pellegrini is in the Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mail: pellegrini@physics.ucla.edu

1. Shintake, T. *et al.* *Nature Photonics* **2**, 555–559 (2008).
2. Arthur, J. *et al.* www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-593.pdf (2002).
3. Altarelli, M. *et al.* (eds) preprint at <http://xfel.desy.de/tdr/tdr/> (2006).
4. Tanaka, T. & Shintake, T. (eds) preprint at www-xfel.spring8.or.jp/SCSSCDR.pdf (2005).
5. Yokoya, K. *et al.* *Proc. 5th Eur. Particle Accelerator Conf. (EPAC96)*, Sitges, Barcelona, KEK preprint 96-68 (1996).
6. Schlenvoigt, H.-P. *et al.* *Nature Phys.* **4**, 130–133 (2008).
7. Schroeder, C. B., Fawley, W. M., Esarey, E. & Leemans, W. P. *Proc. FEL 2006*, preprint at www.bessy.de/fel2006/proceedings/PAPERS/TUPPH055.PDF (2006).
8. Rosenzweig, J. B. *et al.* *Nucl. Instrum. Methods A* **593**, 39–44 (2008).

OPTICS

Farewell to Flatland

Ortwin Hess

Metamaterials are the key to perfect lenses, 'invisibility' cloaks and slow and stored broadband light. A three-dimensional optical metamaterial with a negative refractive index has now been created.

Picture a straw in a glass of water. Looking at the glass from the side, the straw seems to break at the water surface and continue inside the liquid slightly shifted to one side and slightly wider than in the air. The difference between the 'optical density' of water and air, which is responsible for this phenomenon, is usually expressed in terms of the materials' refractive indices — that of water is higher than that of air. All naturally occurring materials, such as water, have a positive refractive index. However, if water had a negative refractive index, the straw would seem to be slightly shifted at the interface but would continue inside the water as if it had been shifted 'the wrong way' (Fig. 1, overleaf).

This unusual light-bending property is a feature of metamaterials — artificially engineered structures — with a negative refractive index. Part of the excitement surrounding these materials is that they could be engineered to 'cloak' objects from electromagnetic radiation such as light — that is, make them seem invisible¹. On page 376 of this issue, Valentine *et al.*² report a three-dimensional 'fishnet' metamaterial consisting of alternating layers of silver and magnesium fluoride. This exceptional material exhibits a negative refractive index across a broad region of the electromagnetic spectrum.

For some time, Victor Veselago's 1960s Gedankenexperiment of a negative refractive index³ seemed little more than just that, a 'thought experiment'. It was not until 2000, when John Pendry proposed that a material

with a negative refractive index would make a 'perfect' lens⁴, that negative refraction began to attract increasing attention⁵. But optical negative-index metamaterials have so far effectively been two-dimensional surfaces, very much at home in the two-dimensional Flatland world of Edwin Abbott's classic novella⁶.

Recently, materials such as semiconductor metamaterials⁷ and composite nanowire materials⁸, which display negative refraction but do not have a bulk negative refractive index, have been manufactured. So how can we tell that the remarkable chunk of nano-structured material reported by Valentine and colleagues² really does exhibit a bulk three-dimensional negative refractive index? To confirm this, we will need to be sure of two features of the material.

First, are the unit cells, the 'atoms' of the metamaterial, sufficiently small relative to the wavelength (colour) of light for the material viewed with such light to appear homogeneous, and for the individual unit cells not to be apparent? The dimensions of the material's unit cell (80 nm × 860 nm × 265 nm) mean that it is not yet quite isotropic. However, particularly in the case of light travelling close to perpendicularly to the fishnet-structure layers, the periodicity (separation) of the layers is clearly much smaller than the light's wavelength. Moreover, the metamaterial's dispersion relation, which describes the way wave propagation varies with wavelength, reveals that in the negative-index regime there is a single propagating mode (with approximately fixed frequency) that can



50 YEARS AGO

The Committee for Biological Acoustics organized a demonstration meeting to illustrate the wide range of research in this comparatively new and expanding field ... The progress made in the United States in underwater recording was emphasized by a demonstration of the acoustic emission of whales by Dr. W. E. Schevill ... Mr E. F. Woods, of the B.B.C., demonstrated his work on the sounds of colonies of honey bees, while developing an electronic device, the 'Apidictor', which can assist apiarists in detecting the onset of swarming in hives.

From *Nature* 20 September 1958.

100 YEARS AGO

Coast Erosion and Foreshore Protection. By John S. Owens and Gerald O. Case — The authors very properly point out that there is no one method of protection that can be applied to all coasts, but that each shore much be considered on its merits, and that it is only after due consideration has been given to the special circumstances which may influence the effect of the sea upon any particular shore that the proper remedy can be designed ... The chapter on ferro-concrete groynes contains much useful informal information on the application of this material to sea defence work, and gives illustrations and cost of works carried out for the protection of the coast of Sussex. The cost of these groynes is given as twenty shillings a foot run, which compares favourably with timber.

ALSO:

An appliance for working the keyboard of a typewriter on a type-setting machine from a distance by means of wireless telegraphy has been devised by Mr. Hans Knudsen, and a demonstration of the experimental apparatus was given at the Hotel Cecil on Thursday last.

From *Nature* 24 September 1908.

50 & 100 YEARS AGO

CLIMATE SCIENCE

A bright side of precipitation

During the period 1986–2000, the world — or at least its terrestrial parts — became wetter. On average, the increase in precipitation over this time was 3.5 mm yr^{-1} . In other words, the hydrological cycle intensified, and Martin Wild and his colleagues set themselves the task of investigating the factors responsible (M. Wild, J. Grieser and C. Schär *Geophys. Res. Lett.* **35**, L17706, doi:10.1029/2008GL034842; 2008).

The strength of the main engine that powers the hydrological cycle depends on the balance of radiative energy at Earth's surface. That balance in turn depends, first, on the amount of solar radiation absorbed by the surface, and second, on the difference between the amount of thermal radiation being transmitted

upwards into the atmosphere (upward radiation) and that being re-radiated back (downward radiation, which is controlled by the greenhouse capacity of the atmosphere). The sum of these fluxes determines the amount of energy that is available for the latent heat of evaporation, which drives precipitation. Any change in these fluxes induces changes in evaporation and precipitation, and thus in the intensity of the hydrological cycle.

For their calculations, the authors drew from two sources: the Global Energy Balance Archive and the Baseline Surface Radiation Network. The latter provides more sophisticated data, but started up only in the 1990s. Another consideration in the calculations was the serious but temporary



S. CURRY/AFP/GETTY IMAGES

effects on radiation balance of the huge emissions from the 1991 eruption of Mount Pinatubo in the Philippines, which lasted for several years.

After accounting for those effects, Wild *et al.* conclude that there has been a steady rise in net surface radiation of $0.2 \text{ W m}^{-2} \text{ yr}^{-1}$ between 1986 and 2000. They mention two possible causes. One is 'solar brightening' — a rise in surface radiation resulting from a more transparent atmosphere caused by a decrease in anthropogenic aerosols. The other is a more

powerful greenhouse effect, increasing downward radiation.

As the authors point out, their study cannot provide a truly global picture because of the lack of data from the oceans. But their calculated rise in net surface radiation fits well with the estimate of energy flux required to drive the rise in precipitation. This provides a satisfying first-order numerical link between the changes in precipitation and radiation, recorded in independent data sets, over the 15-year period of their study.

Tim Lincoln

collaboration (Fig. 1). For all their promise, the LCLS and XFEL are large and expensive; the LCLS is about 2 kilometres long, and XFEL will be even longer. The construction cost for the LCLS, which uses an existing particle accelerator, is about €300 million (\$450 million), and XFEL will cost about €1 billion.

Shintake and colleagues are part of a team designing a Japanese X-FEL called the SPring-8 Compact Self-Amplified Spontaneous Emission Source (SCSS), which is planned for construction in the near future⁴. The work they now describe¹ uses a smaller-scale FEL, working at UV wavelengths, to show that cheaper and smaller X-FELs can be made by taking advantage of advances in the physics and technology of electron beams.

FELs operate by combining the technology of lasers and electron accelerators (Fig. 2). A particle accelerator produces electrons moving at close to the speed of light. These pass through a magnet, which causes them to follow

a sinusoidal trajectory. The acceleration due to the bending of the electrons by this 'undulator' magnet produces electromagnetic waves. These have a wavelength that is proportional to the oscillation period of the sinusoidal path and inversely proportional to the square of the electron energy. Typically, the oscillation has a period of a few centimetres, and therefore, to produce X-rays with a wavelength of about 10^{-10} metres (1 ångström, the typical size of an atom), electrons are required that have energies of up to 15 gigaelectronvolts. For comparison, the energy of electrons produced by medical electron linear accelerators — used in many hospitals for radiation therapy — is 1,000 times smaller than this value. Medical accelerators are about a metre long; if an X-FEL accelerator were to be constructed using the same technology, it would need to be around a kilometre long. The length of the undulator magnet alone would be about 100 metres, and so, including space for X-ray diagnostic and manipulation

equipment, the entire X-FEL would measure a few kilometres.

One way to make X-FEL sources more manageable would be to reduce the electron-beam energy but to force the electrons into tighter undulations. This would have the effect of making both the accelerator and undulator shorter. Shintake *et al.*¹ show that the accelerator length can be reduced by using a 'C-band' high-frequency accelerator, a technology developed for the construction of high-energy electron–positron linear colliders⁵. Other advances involved are an improved electron source and an undulator magnet with smaller period and shorter length. When combined in the SCSS design, these advances would lead to a more compact X-FEL.

An even more advanced instrument under consideration^{6,7} is a high-gradient accelerator driven by a phenomenon known as a laser-plasma wakefield, which is created by passing a brief laser pulse through a plasma of charged particles. Whereas Shintake and colleagues' approach can decrease the length of the accelerator to around a half or a third of that possible now, the laser-plasma accelerator could provide a 100-fold reduction. However, it is still at the exploration and proof-of-principle stage.

The size and cost of an X-FEL could also be reduced by using sources that produce electron beams with higher electron density and smaller variation in momentum than are currently available; together, these properties are described as the beam emittance. Smaller-emittance electron beams would allow the use of lower electron energies, and shortened accelerators and undulators. Emittance could

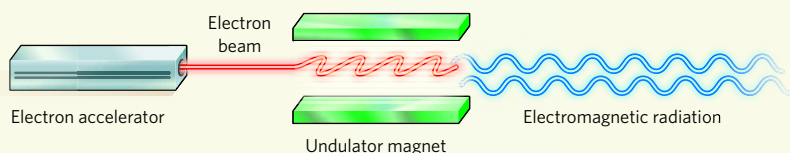


Figure 2 | X-ray specs. A free-electron laser (FEL) consists of a linear accelerator producing a beam of high-energy electrons that then pass through an 'undulator' magnet. This magnet bends the electrons into a sinusoidal path, causing them to emit electromagnetic radiation whose frequency depends on the energy of the electrons and the wavelength of their undulations. Using current technology, a FEL that emits X-rays would be several kilometres long, but the approach tested by Shintake *et al.*¹ on an ultraviolet-emitting device could reduce the size two- to threefold.

be reduced by optimizing the charge and volume of the bunches of electrons produced from the source⁸, a prediction confirmed by initial results at SLAC (P. Emma and D. Dowell, personal communication). Combining this approach with a high-gradient accelerator such as that used by Shintake *et al.*¹ would reduce the size and cost of FELs still further.

Because any reduction in cost and size also implies a reduction in the average number of photons produced, some experiments will always require large facilities such as the LCLS or XFEL. The new smaller systems would nevertheless add significantly to our experimental capabilities. There is a new world to explore at atomic and molecular scales of time and distance, and a multiplicity of approaches, such as that described by Shintake *et al.*, together with other FEL

systems, would speed up the exploration of this uncharted territory. ■

Claudio Pellegrini is in the Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mail: pellegrini@physics.ucla.edu

1. Shintake, T. *et al.* *Nature Photonics* **2**, 555–559 (2008).
2. Arthur, J. *et al.* www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-593.pdf (2002).
3. Altarelli, M. *et al.* (eds) preprint at <http://xfel.desy.de/tdr/tdr/> (2006).
4. Tanaka, T. & Shintake, T. (eds) preprint at www-xfel.spring8.or.jp/SCSSCDR.pdf (2005).
5. Yokoya, K. *et al.* *Proc. 5th Eur. Particle Accelerator Conf. (EPAC96)*, Sitges, Barcelona, KEK preprint 96-68 (1996).
6. Schlenvoigt, H.-P. *et al.* *Nature Phys.* **4**, 130–133 (2008).
7. Schroeder, C. B., Fawley, W. M., Esarey, E. & Leemans, W. P. *Proc. FEL 2006*, preprint at www.bessy.de/fel2006/proceedings/PAPERS/TUPPH055.PDF (2006).
8. Rosenzweig, J. B. *et al.* *Nucl. Instrum. Methods A* **593**, 39–44 (2008).

OPTICS

Farewell to Flatland

Ortwin Hess

Metamaterials are the key to perfect lenses, 'invisibility' cloaks and slow and stored broadband light. A three-dimensional optical metamaterial with a negative refractive index has now been created.

Picture a straw in a glass of water. Looking at the glass from the side, the straw seems to break at the water surface and continue inside the liquid slightly shifted to one side and slightly wider than in the air. The difference between the 'optical density' of water and air, which is responsible for this phenomenon, is usually expressed in terms of the materials' refractive indices — that of water is higher than that of air. All naturally occurring materials, such as water, have a positive refractive index. However, if water had a negative refractive index, the straw would seem to be slightly shifted at the interface but would continue inside the water as if it had been shifted 'the wrong way' (Fig. 1, overleaf).

This unusual light-bending property is a feature of metamaterials — artificially engineered structures — with a negative refractive index. Part of the excitement surrounding these materials is that they could be engineered to 'cloak' objects from electromagnetic radiation such as light — that is, make them seem invisible¹. On page 376 of this issue, Valentine *et al.*² report a three-dimensional 'fishnet' metamaterial consisting of alternating layers of silver and magnesium fluoride. This exceptional material exhibits a negative refractive index across a broad region of the electromagnetic spectrum.

For some time, Victor Veselago's 1960s Gedankenexperiment of a negative refractive index³ seemed little more than just that, a 'thought experiment'. It was not until 2000, when John Pendry proposed that a material

with a negative refractive index would make a 'perfect' lens⁴, that negative refraction began to attract increasing attention⁵. But optical negative-index metamaterials have so far effectively been two-dimensional surfaces, very much at home in the two-dimensional Flatland world of Edwin Abbott's classic novella⁶.

Recently, materials such as semiconductor metamaterials⁷ and composite nanowire materials⁸, which display negative refraction but do not have a bulk negative refractive index, have been manufactured. So how can we tell that the remarkable chunk of nano-structured material reported by Valentine and colleagues² really does exhibit a bulk three-dimensional negative refractive index? To confirm this, we will need to be sure of two features of the material.

First, are the unit cells, the 'atoms' of the metamaterial, sufficiently small relative to the wavelength (colour) of light for the material viewed with such light to appear homogeneous, and for the individual unit cells not to be apparent? The dimensions of the material's unit cell (80 nm × 860 nm × 265 nm) mean that it is not yet quite isotropic. However, particularly in the case of light travelling close to perpendicularly to the fishnet-structure layers, the periodicity (separation) of the layers is clearly much smaller than the light's wavelength. Moreover, the metamaterial's dispersion relation, which describes the way wave propagation varies with wavelength, reveals that in the negative-index regime there is a single propagating mode (with approximately fixed frequency) that can



50 YEARS AGO

The Committee for Biological Acoustics organized a demonstration meeting to illustrate the wide range of research in this comparatively new and expanding field ... The progress made in the United States in underwater recording was emphasized by a demonstration of the acoustic emission of whales by Dr. W. E. Schevill ... Mr E. F. Woods, of the B.B.C., demonstrated his work on the sounds of colonies of honey bees, while developing an electronic device, the 'Apidictor', which can assist apiarists in detecting the onset of swarming in hives.

From *Nature* 20 September 1958.

100 YEARS AGO

Coast Erosion and Foreshore Protection. By John S. Owens and Gerald O. Case — The authors very properly point out that there is no one method of protection that can be applied to all coasts, but that each shore much be considered on its merits, and that it is only after due consideration has been given to the special circumstances which may influence the effect of the sea upon any particular shore that the proper remedy can be designed ... The chapter on ferro-concrete groynes contains much useful informal information on the application of this material to sea defence work, and gives illustrations and cost of works carried out for the protection of the coast of Sussex. The cost of these groynes is given as twenty shillings a foot run, which compares favourably with timber.

ALSO:

An appliance for working the keyboard of a typewriter on a type-setting machine from a distance by means of wireless telegraphy has been devised by Mr. Hans Knudsen, and a demonstration of the experimental apparatus was given at the Hotel Cecil on Thursday last.

From *Nature* 24 September 1908.

50 & 100 YEARS AGO

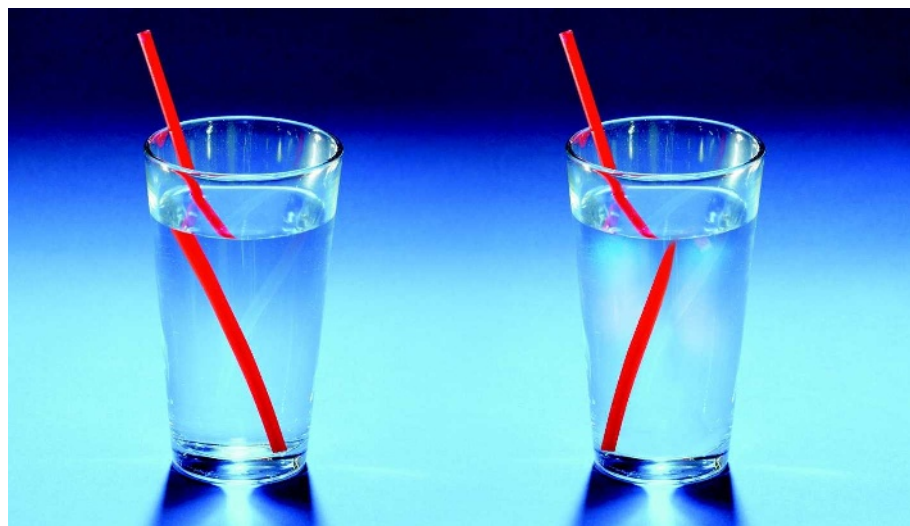


Figure 1 | Negative refraction. The difference in the optical density of air and 'normal' water (left) causes a straw in a glass of water to seem to be shifted at the interface and slightly enlarged inside the liquid. In 'negative-index water' (right), the straw would seem to continue in 'the wrong direction'. This is the principle that applies to negative-index metamaterials, such as that created by Valentine *et al.*².

be associated with a propagation constant and, in turn, with an effective refractive index.

Second, does the material exhibit a negative phase velocity? In other words, how can we tell that the phase of a forward-moving wave is moving backwards inside the material? This might be possible if, to measure a beam deflected in the 'wrong' direction, we apply Snell's law⁹, which relates the angles of incidence and refraction of a wave. However, even everyday optical materials, such as a bathroom mirror, may at least partially display the effects of negative refraction yet have a positive refractive index¹⁰. Indeed, it is now established that negative refraction can occur in materials displaying anisotropy, such as calcite, and in two- or three-dimensionally structured materials such as photonic crystals¹¹.

To identify the signatures of negative refraction from a negative refractive index, Valentine *et al.*² combined experimental results with theoretical modelling. Experimentally, the authors did turn to Snell's law to determine the refractive index, but crucially they retrieved its value from the far-field transmission pattern of a prism configuration representing a signature of the (backward) phase evolution inside the metamaterial. Together with supporting theoretical simulations that show a backward-moving phase inside the prism, this strongly indicates that Valentine and colleagues' fishnet structure is indeed a three-dimensional material with a negative refractive index.

Another point in considering Valentine and colleagues' material is whether, when light is shone onto it, the light lost passing through it is low. To be useful, materials should certainly not absorb all light. However, because the constituent elements of a metamaterial, particularly metals, always exhibit a loss in some part of the frequency spectrum, the engineering of losses is an important task. A promising solution is to use nano-structuring (tightly coupled

nano-plasmonic resonators) to simultaneously increase the bandwidth and effectively shift the losses to a frequency range where the optics is not impaired. In this respect, the metamaterial described by Valentine *et al.*² scores very well, with a record-breaking 'figure of merit' of 3.5, indicating low loss. In addition, this metamaterial displays a negative refractive index over a remarkably wide window of the spectrum, ranging from around 1,800 nm to less than 1,500 nm.

Finally, what uses might there be for a broadband, low-loss, three-dimensional optical

material with a negative refractive index? First, it would be a way of constructing Pendry's perfect lens⁴, and thus imaging a three-dimensional object the size of a molecule. Second, it would be beneficial for electromagnetic cloaking¹ and its applications, in that a cloaking material that absorbs less radiation will ideally prevent the cloaked object from appearing as a dark silhouette. Third, the ability to slow down and store broadband light using the 'trapped rainbow' principle¹² has clearly come closer to reality.

Saying goodbye to Flatland and entering Spaceland was not only a great achievement for A. Square, the protagonist in Abbott's novella. It is also a promising step towards producing optical negative-index metamaterials. The prospects that this step offers for research involving the control and manipulation of light could hardly be more exciting.

Ortwin Hess is at the Advanced Technology Institute and Department of Physics, University of Surrey, Guildford GU2 7XH, UK.

e-mail: o.hess@surrey.ac.uk

1. Pendry, J. B., Shurig, D. & Smith, D. R. *Science* **312**, 1780–1782 (2006).
2. Valentine, J. *et al.* *Nature* **455**, 376–379 (2008).
3. Veselago, V. G. *Sov. Phys. Usp.* **10**, 509–514 (1968).
4. Pendry, J. B. *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
5. Shalaev, V. M. *Nature Photonics* **1**, 41–48 (2007).
6. Abbott, E. A. *Flatland* (Seeley, 1884).
7. Hoffman, A. J. *et al.* *Nature Mater.* **6**, 946–950 (2007).
8. Yao, J. *et al.* *Science* **321**, 930 (2008).
9. Born, M. & Wolf, E. *Principles of Optics* (Cambridge Univ. Press, 1999).
10. Wegener, M., Dolling, G. & Linden, S. *Nature Mater.* **6**, 475–476 (2007).
11. Cubukcu, E. *et al.* *Nature* **423**, 604–605 (2003).
12. Tsakmakidis, K. L., Boardman, A. D. & Hess, O. *Nature* **450**, 397–401 (2007).

BIOCHEMISTRY

Divergence from the superfamily

Lawrence J. Marnett

The unique structure of an enzyme that synthesizes lipid mediators in plants defines its function and serves as a characteristic footprint in genomes of plants and some animals.

In plants and animals, molecular signalling pathways that are mediated by metabolites of polyunsaturated fatty acids are essential for communication at cellular and organismal levels. For example, the plant metabolites known as jasmonates and green-leaf volatiles control not only growth but also responses to wounding and infection^{1,2}, and contribute to the characteristic taste and smell of certain plants. Jasmonates and green-leaf volatiles are formed from fatty-acid hydroperoxides through the action of the allene oxide synthase (AOS) and hydroperoxide lyase (HPL) enzymes, respectively^{3,4}. Exactly how these enzymes react with their hydroperoxide substrates has remained unclear. The crystal structures of AOS — as

free enzyme and in complex with substrate and intermediate analogues — are presented by Lee *et al.*⁵ on page 363 of this issue, and provide fascinating clues.

Both AOSs and HPLs are members of the CYP74 family, which belongs to the cytochrome P450 (CYP) superfamily. CYP proteins contain a haem group, and are defined by the characteristic absorption of their ferrous carbon monoxide complex at a wavelength of 450 nanometres. These enzymes catalyse oxidation reactions that are central to many metabolic pathways in a wide range of organisms — for example, they mediate steroid biosynthesis in humans and nutrient breakdown in bacteria⁶. To do this, they take two electrons from a reduct-

be reduced by optimizing the charge and volume of the bunches of electrons produced from the source⁸, a prediction confirmed by initial results at SLAC (P. Emma and D. Dowell, personal communication). Combining this approach with a high-gradient accelerator such as that used by Shintake *et al.*¹ would reduce the size and cost of FELs still further.

Because any reduction in cost and size also implies a reduction in the average number of photons produced, some experiments will always require large facilities such as the LCLS or XFEL. The new smaller systems would nevertheless add significantly to our experimental capabilities. There is a new world to explore at atomic and molecular scales of time and distance, and a multiplicity of approaches, such as that described by Shintake *et al.*, together with other FEL

systems, would speed up the exploration of this uncharted territory. ■

Claudio Pellegrini is in the Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095, USA.
e-mail: pellegrini@physics.ucla.edu

1. Shintake, T. *et al.* *Nature Photonics* **2**, 555–559 (2008).
2. Arthur, J. *et al.* www.slac.stanford.edu/cgi-wrap/getdoc/slac-r-593.pdf (2002).
3. Altarelli, M. *et al.* (eds) preprint at <http://xfel.desy.de/tdr/tdr/> (2006).
4. Tanaka, T. & Shintake, T. (eds) preprint at www-xfel.spring8.or.jp/SCSSCDR.pdf (2005).
5. Yokoya, K. *et al.* *Proc. 5th Eur. Particle Accelerator Conf. (EPAC96)*, Sitges, Barcelona, KEK preprint 96-68 (1996).
6. Schlenvoigt, H.-P. *et al.* *Nature Phys.* **4**, 130–133 (2008).
7. Schroeder, C. B., Fawley, W. M., Esarey, E. & Leemans, W. P. *Proc. FEL 2006*, preprint at www.bessy.de/fel2006/proceedings/PAPERS/TUPPH055.PDF (2006).
8. Rosenzweig, J. B. *et al.* *Nucl. Instrum. Methods A* **593**, 39–44 (2008).

OPTICS

Farewell to Flatland

Ortwin Hess

Metamaterials are the key to perfect lenses, 'invisibility' cloaks and slow and stored broadband light. A three-dimensional optical metamaterial with a negative refractive index has now been created.

Picture a straw in a glass of water. Looking at the glass from the side, the straw seems to break at the water surface and continue inside the liquid slightly shifted to one side and slightly wider than in the air. The difference between the 'optical density' of water and air, which is responsible for this phenomenon, is usually expressed in terms of the materials' refractive indices — that of water is higher than that of air. All naturally occurring materials, such as water, have a positive refractive index. However, if water had a negative refractive index, the straw would seem to be slightly shifted at the interface but would continue inside the water as if it had been shifted 'the wrong way' (Fig. 1, overleaf).

This unusual light-bending property is a feature of metamaterials — artificially engineered structures — with a negative refractive index. Part of the excitement surrounding these materials is that they could be engineered to 'cloak' objects from electromagnetic radiation such as light — that is, make them seem invisible¹. On page 376 of this issue, Valentine *et al.*² report a three-dimensional 'fishnet' metamaterial consisting of alternating layers of silver and magnesium fluoride. This exceptional material exhibits a negative refractive index across a broad region of the electromagnetic spectrum.

For some time, Victor Veselago's 1960s Gedankenexperiment of a negative refractive index³ seemed little more than just that, a 'thought experiment'. It was not until 2000, when John Pendry proposed that a material

with a negative refractive index would make a 'perfect' lens⁴, that negative refraction began to attract increasing attention⁵. But optical negative-index metamaterials have so far effectively been two-dimensional surfaces, very much at home in the two-dimensional Flatland world of Edwin Abbott's classic novella⁶.

Recently, materials such as semiconductor metamaterials⁷ and composite nanowire materials⁸, which display negative refraction but do not have a bulk negative refractive index, have been manufactured. So how can we tell that the remarkable chunk of nano-structured material reported by Valentine and colleagues² really does exhibit a bulk three-dimensional negative refractive index? To confirm this, we will need to be sure of two features of the material.

First, are the unit cells, the 'atoms' of the metamaterial, sufficiently small relative to the wavelength (colour) of light for the material viewed with such light to appear homogeneous, and for the individual unit cells not to be apparent? The dimensions of the material's unit cell (80 nm × 860 nm × 265 nm) mean that it is not yet quite isotropic. However, particularly in the case of light travelling close to perpendicularly to the fishnet-structure layers, the periodicity (separation) of the layers is clearly much smaller than the light's wavelength. Moreover, the metamaterial's dispersion relation, which describes the way wave propagation varies with wavelength, reveals that in the negative-index regime there is a single propagating mode (with approximately fixed frequency) that can



50 YEARS AGO

The Committee for Biological Acoustics organized a demonstration meeting to illustrate the wide range of research in this comparatively new and expanding field ... The progress made in the United States in underwater recording was emphasized by a demonstration of the acoustic emission of whales by Dr. W. E. Schevill ... Mr E. F. Woods, of the B.B.C., demonstrated his work on the sounds of colonies of honey bees, while developing an electronic device, the 'Apidictor', which can assist apiarists in detecting the onset of swarming in hives.

From *Nature* 20 September 1958.

100 YEARS AGO

Coast Erosion and Foreshore Protection. By John S. Owens and Gerald O. Case — The authors very properly point out that there is no one method of protection that can be applied to all coasts, but that each shore much be considered on its merits, and that it is only after due consideration has been given to the special circumstances which may influence the effect of the sea upon any particular shore that the proper remedy can be designed ... The chapter on ferro-concrete groynes contains much useful informal information on the application of this material to sea defence work, and gives illustrations and cost of works carried out for the protection of the coast of Sussex. The cost of these groynes is given as twenty shillings a foot run, which compares favourably with timber.

ALSO:

An appliance for working the keyboard of a typewriter on a type-setting machine from a distance by means of wireless telegraphy has been devised by Mr. Hans Knudsen, and a demonstration of the experimental apparatus was given at the Hotel Cecil on Thursday last.

From *Nature* 24 September 1908.

50 & 100 YEARS AGO

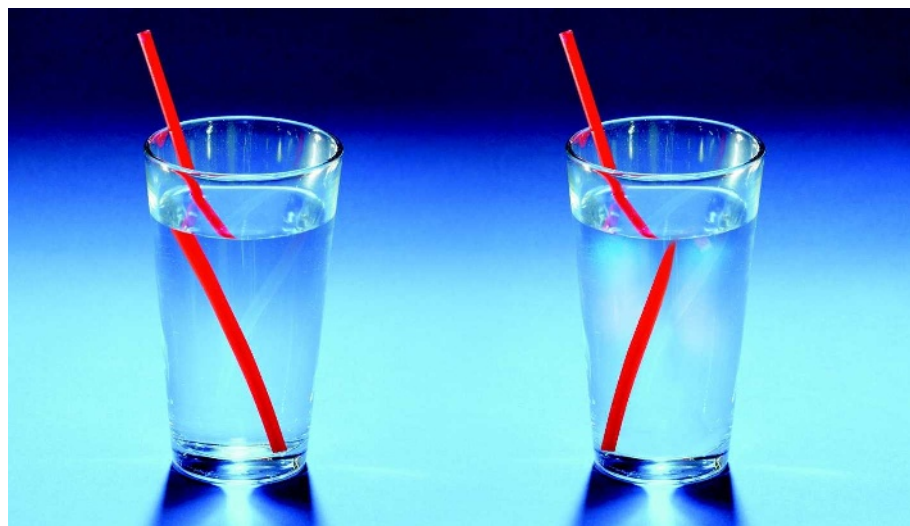


Figure 1 | Negative refraction. The difference in the optical density of air and 'normal' water (left) causes a straw in a glass of water to seem to be shifted at the interface and slightly enlarged inside the liquid. In 'negative-index water' (right), the straw would seem to continue in 'the wrong direction'. This is the principle that applies to negative-index metamaterials, such as that created by Valentine *et al.*².

be associated with a propagation constant and, in turn, with an effective refractive index.

Second, does the material exhibit a negative phase velocity? In other words, how can we tell that the phase of a forward-moving wave is moving backwards inside the material? This might be possible if, to measure a beam deflected in the 'wrong' direction, we apply Snell's law⁹, which relates the angles of incidence and refraction of a wave. However, even everyday optical materials, such as a bathroom mirror, may at least partially display the effects of negative refraction yet have a positive refractive index¹⁰. Indeed, it is now established that negative refraction can occur in materials displaying anisotropy, such as calcite, and in two- or three-dimensionally structured materials such as photonic crystals¹¹.

To identify the signatures of negative refraction from a negative refractive index, Valentine *et al.*² combined experimental results with theoretical modelling. Experimentally, the authors did turn to Snell's law to determine the refractive index, but crucially they retrieved its value from the far-field transmission pattern of a prism configuration representing a signature of the (backward) phase evolution inside the metamaterial. Together with supporting theoretical simulations that show a backward-moving phase inside the prism, this strongly indicates that Valentine and colleagues' fishnet structure is indeed a three-dimensional material with a negative refractive index.

Another point in considering Valentine and colleagues' material is whether, when light is shone onto it, the light lost passing through it is low. To be useful, materials should certainly not absorb all light. However, because the constituent elements of a metamaterial, particularly metals, always exhibit a loss in some part of the frequency spectrum, the engineering of losses is an important task. A promising solution is to use nano-structuring (tightly coupled

nano-plasmonic resonators) to simultaneously increase the bandwidth and effectively shift the losses to a frequency range where the optics is not impaired. In this respect, the metamaterial described by Valentine *et al.*² scores very well, with a record-breaking 'figure of merit' of 3.5, indicating low loss. In addition, this metamaterial displays a negative refractive index over a remarkably wide window of the spectrum, ranging from around 1,800 nm to less than 1,500 nm.

Finally, what uses might there be for a broadband, low-loss, three-dimensional optical

material with a negative refractive index? First, it would be a way of constructing Pendry's perfect lens⁴, and thus imaging a three-dimensional object the size of a molecule. Second, it would be beneficial for electromagnetic cloaking¹ and its applications, in that a cloaking material that absorbs less radiation will ideally prevent the cloaked object from appearing as a dark silhouette. Third, the ability to slow down and store broadband light using the 'trapped rainbow' principle¹² has clearly come closer to reality.

Saying goodbye to Flatland and entering Spaceland was not only a great achievement for A. Square, the protagonist in Abbott's novella. It is also a promising step towards producing optical negative-index metamaterials. The prospects that this step offers for research involving the control and manipulation of light could hardly be more exciting.

Ortwin Hess is at the Advanced Technology Institute and Department of Physics, University of Surrey, Guildford GU2 7XH, UK.

e-mail: o.hess@surrey.ac.uk

1. Pendry, J. B., Shurig, D. & Smith, D. R. *Science* **312**, 1780–1782 (2006).
2. Valentine, J. *et al.* *Nature* **455**, 376–379 (2008).
3. Veselago, V. G. *Sov. Phys. Usp.* **10**, 509–514 (1968).
4. Pendry, J. B. *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
5. Shalaev, V. M. *Nature Photonics* **1**, 41–48 (2007).
6. Abbott, E. A. *Flatland* (Seeley, 1884).
7. Hoffman, A. J. *et al.* *Nature Mater.* **6**, 946–950 (2007).
8. Yao, J. *et al.* *Science* **321**, 930 (2008).
9. Born, M. & Wolf, E. *Principles of Optics* (Cambridge Univ. Press, 1999).
10. Wegener, M., Dolling, G. & Linden, S. *Nature Mater.* **6**, 475–476 (2007).
11. Cubukcu, E. *et al.* *Nature* **423**, 604–605 (2003).
12. Tsakmakidis, K. L., Boardman, A. D. & Hess, O. *Nature* **450**, 397–401 (2007).

BIOCHEMISTRY

Divergence from the superfamily

Lawrence J. Marnett

The unique structure of an enzyme that synthesizes lipid mediators in plants defines its function and serves as a characteristic footprint in genomes of plants and some animals.

In plants and animals, molecular signalling pathways that are mediated by metabolites of polyunsaturated fatty acids are essential for communication at cellular and organismal levels. For example, the plant metabolites known as jasmonates and green-leaf volatiles control not only growth but also responses to wounding and infection^{1,2}, and contribute to the characteristic taste and smell of certain plants. Jasmonates and green-leaf volatiles are formed from fatty-acid hydroperoxides through the action of the allene oxide synthase (AOS) and hydroperoxide lyase (HPL) enzymes, respectively^{3,4}. Exactly how these enzymes react with their hydroperoxide substrates has remained unclear. The crystal structures of AOS — as

free enzyme and in complex with substrate and intermediate analogues — are presented by Lee *et al.*⁵ on page 363 of this issue, and provide fascinating clues.

Both AOSs and HPLs are members of the CYP74 family, which belongs to the cytochrome P450 (CYP) superfamily. CYP proteins contain a haem group, and are defined by the characteristic absorption of their ferrous carbon monoxide complex at a wavelength of 450 nanometres. These enzymes catalyse oxidation reactions that are central to many metabolic pathways in a wide range of organisms — for example, they mediate steroid biosynthesis in humans and nutrient breakdown in bacteria⁶. To do this, they take two electrons from a reduct-

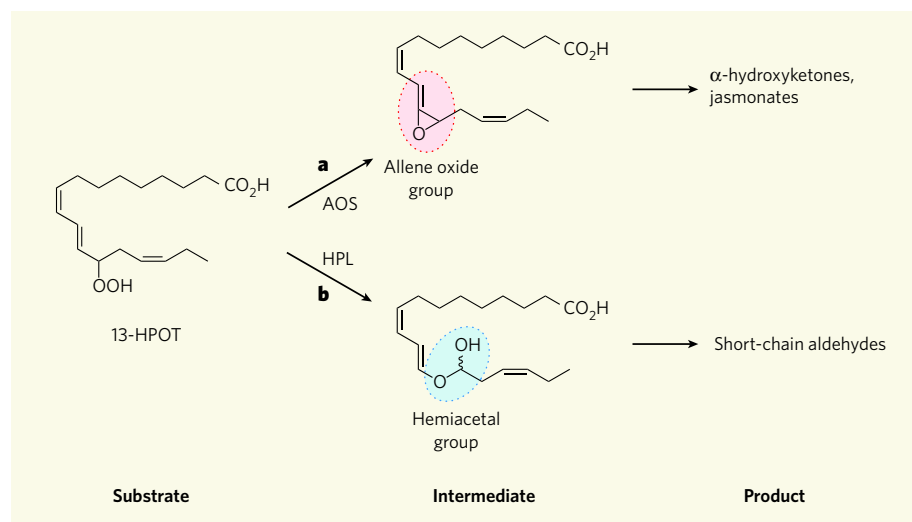


Figure 1 | CYP74 enzyme function. A partially oxidized lipid called 13-hydroperoxyoctadeca-9,11,15-trienoic acid (13-HPOT) is a typical substrate for the plant CYP74 enzymes allene oxide synthase (AOS) and hydroperoxide lyase (HPL). **a**, AOS converts this substrate to an unstable allene oxide, which then hydrolyses to α -hydroxyketones and jasmonates. **b**, Similarly, HPL converts 13-HPOT to a hemiacetal intermediate, which is then broken into short-chain aldehydes such as green-leaf volatiles. Lee *et al.*⁵ solve the structure of a plant AOS to show which structural features are responsible for the unique function of these cytochrome P450 enzymes, compared with other members of this superfamily.

ase enzyme, which itself gets the electrons from another molecule. CYPs then bind to an oxygen molecule, splitting it into water and an oxygen atom that the enzyme eventually transfers to a substrate.

Members of the CYP74 family are atypical in that they use fatty-acid hydroperoxides as a source of the oxygen atom to be transferred, which eliminates the need for an external source of electrons or a reductase⁷. What's more, they use peroxide substrates to form new metabolites^{8,9}. For example, AOS from the plant *Arabidopsis thaliana* (At-AOS) and the related HPL enzyme require neither an external source of electrons nor association with a reductase. Instead, in a multi-step reaction, they transform a partially oxidized lipid such as 13-hydroperoxyoctadeca-9,11,15-trienoic acid (13-HPOT) to jasmonates and green-leaf volatiles, respectively (Fig. 1).

Many CYPs are membrane-bound, making it difficult to crystallize them for structural analysis. Nevertheless, nearly 30 CYP structures have been solved, all of which display a characteristic globular fold that mainly contains α -helices¹⁰. A long helix (the I-helix) runs through the centre of the protein and borders the distal face of the haem group. The shorter L-helix borders the proximal face of the haem, which consists of the haem-binding loop. A cysteine residue within this loop binds to the iron in the haem. The oxidized substrates bind in a cavity on the distal side of the haem, and electrons are transferred following association of a reductase near the proximal side of the haem.

Lee *et al.*⁵ describe the first structure of a plant CYP, that of At-AOS. They find that certain unique structural elements within this enzyme set its catalytic activity apart from that of the known structures of other CYPs. Of particular

interest is an insertion of nine amino-acid residues in the middle of the haem-binding loop of At-AOS, which distorts the loop and alters the putative CYP-reductase interface. The insertion is conserved in — and seems to be unique to — all CYP74 family members, because none of the 8,000 or so other CYPs contains an insertion of even a single amino acid in the haem-binding loop. Moreover, the authors find that alterations of amino-acid residues in the I-helix result in two changes in enzyme topology. First, there is an asparagine residue directly over the distal side of the haem, which blocks oxygen binding. Second, a crucial threonine hydroxyl group that is necessary for typical CYP oxygenase function is missing. Thus, At-AOS cannot function as an oxygenase but instead is optimized for hydroperoxide rearrangement.

Conversion of hydroperoxide to allene oxide involves formation of an epoxyallylic radical intermediate. Lee and colleagues' structures of At-AOS in complex with substrate and intermediate analogues suggest that the eleventh carbon atom (C-11) of this intermediate — where the radical forms — is positioned within the At-AOS active site, near phenylalanine residue 137. The C-11 radical, and the cation that forms after its oxidation, may be stabilized by interactions between the radical or the cation and the aromatic ring of this residue (see Fig. 2a on page 364).

When phenylalanine-137 is mutated to a leucine residue, At-AOS mediates fragmentation of 13-HPOT to short-chain aldehydes, rather than generating allene oxide from this lipid. Interestingly, just such a phenylalanine-to-leucine substitution occurs naturally in HPLs. These enzymes are designed to catalyse multiple transformations, the partitioning of which is determined by one or two main

residues in their active site. So when the authors mutated the corresponding leucine of an HPL to phenylalanine, they observed some AOS activity, even though HPL did not completely convert to an AOS.

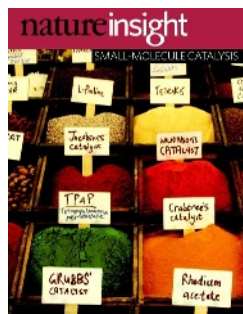
How widespread are the members of the CYP74 family outside the plant kingdom? Using signature sequences from At-AOS, Lee *et al.*⁵ searched genome databases to identify several candidate genes in marine invertebrates and two in rhizobacteria — microorganisms that live symbiotically with plants. Sequence identity between CYP74 candidate genes and their plant counterparts is less than the 40% threshold that is required for them to be classified as a CYP74 family member. Nonetheless, all the candidate proteins contain the distinctive nine-residue insert in their haem-binding loop.

Studying the activity of CYP74-like proteins from bacteria, coral and the primitive marine animal amphioxus, the authors find that these enzymes do catalyse the breakdown of fatty-acid hydroperoxides, although not always to allene oxides. The number of genes that encode these CYP74-like proteins varied between species, and in amphioxus there are as many as 20 such genes. Unlike their plant counterparts, animal CYP74 proteins contain introns — gene sequences that do not contribute to the protein sequence.

The fact that CYP74 proteins are found in many plants, in the last common ancestors of plants and animals, and in less complex animals but not in vertebrates suggests that they reached a dead end early in animal evolution. Perhaps the functions of these enzymes or their metabolic products in vertebrates have been subsumed by other enzymes — such as AOSs from another, catalase-like enzyme family¹¹ — or fatty-acid derivatives that evolved as Earth's atmosphere became more aerobic. In animals, for example, prostaglandin endoperoxides are converted by non-oxygenating CYPs to thromboxane and prostacyclin — mediators of vascular biology. It will be interesting to see how the mechanisms of these transformations conform to the principles defined for CYP74.

Lawrence J. Marnett is in the Department of Biochemistry, Vanderbilt University, School of Medicine, Nashville, Tennessee 37232-0146, USA. e-mail: larry.marnett@vanderbilt.edu

- Creelman, R. A. & Mullet, J. E. *Proc. Natl Acad. Sci. USA* **92**, 4114–4119 (1995).
- Matsui, K. *Curr. Opin. Plant Biol.* **9**, 274–280 (2006).
- Tijet, N. & Brash, A. R. *Prostaglandins Other Lipid Mediat.* **68–69**, 423–431 (2002).
- Grechkin, A. N. *Prostaglandins Other Lipid Mediat.* **68–69**, 457–470 (2002).
- Lee, D.-S., Nioche, P., Hamberg, M. & Raman, C. S. *Nature* **455**, 363–368 (2008).
- Ortiz de Montellano, P. *Cytochrome P450: Structure, Mechanism, and Biochemistry*, 3rd edn (Plenum, 2005).
- Hryciay, E. G. & O'Brien, P. J. *Arch. Biochem. Biophys.* **147**, 14–27 (1971).
- Song, W. C. & Brash, A. R. *Science* **253**, 781–784 (1991).
- Hecker, M. & Ullrich, V. *J. Biol. Chem.* **264**, 141–150 (1989).
- Johnson, E. F. & Stout, C. D. *Biochem. Biophys. Res. Commun.* **338**, 331–336 (2005).
- Oldham, M. L., Brash, A. R. & Newcomer, M. E. *Proc. Natl Acad. Sci. USA* **102**, 297–302 (2005).

**Cover illustration**

(Courtesy of P. Syder/
photolibary.com.
Artwork by N. Spencer)

Editor, Nature

Philip Campbell

Publishing

Nick Campbell
Samia Burridge
Claudia Banks

Insights Editor

Karl Ziemelis

Production Editor

Davina Dadley-Moore

Senior Art Editor

Martin Harrison

Art Editor

Nik Spencer

Sponsorship

Amélie Pequignot

Production

Jocelyn Hilton

Marketing

Elena Woodstock
Emily Elkins

Editorial Assistant

Nichola O'Brien

SMALL-MOLECULE CATALYSIS

In a world without catalysts, many chemical processes would not occur — at least, not at a rate that has any practical application. Catalysts have thus become indispensable in a wide range of industrial reactions. But the study of these systems can offer much more: it can change our understanding of fundamental chemical concepts and challenge us to rethink the 'rules' of the chemical world. Research in this area moves at breathtaking speed: many catalytic reactions now considered to be 'standard issue' by organic chemists were almost unthinkable just 10 years ago.

The catalytic systems being investigated today range from nanoparticles to transition metals to enzymes, and it would be difficult (if not impossible) to do justice to such diverse research in just a short series of articles. As a result, this Insight focuses on a key sector within the field: what we are calling 'small-molecule' catalysis. This includes transition-metal complexes — which are indeed small compared with proteins or nanoparticles — as well as organic molecules.

The theme is broader than that, though, as it also encompasses the small molecules that are produced in the catalysed reactions. The ability to synthesize and selectively modify small molecules is crucial for many applications, including drug discovery and the search for new materials. The development of new catalysts often makes it possible to generate previously unattainable compounds, which could have unique physical, chemical or biological properties.

But the basic science is as important as the applications: studies of catalytic mechanisms often uncover new modes of chemical reactivity, such as the way heteroatoms affect cross-coupling reactions, forcing us to think anew about the way molecules interact and react. We hope that this Insight will inspire both chemists and non-chemists, by providing fresh ideas about catalysis and the possibilities that it offers for the future.

Andrew Mitchinson, Senior Editor, News & Views
Joshua Finkelstein, Senior Editor

COMMENTARIES

304 The advent and development of organocatalysis

D. W. C. MacMillan

309 Computational prediction of small-molecule catalysts

K. N. Houk & P. H.-Y. Cheong

REVIEWS

314 Carbon-heteroatom bond formation catalysed by organometallic complexes

J. F. Hartwig

323 Natural products as inspiration for the development of asymmetric catalysis

J. T. Mohr, M. R. Krout & B. M. Stoltz

333 Biologically inspired oxidation catalysis

L. Que Jr & W. B. Tolman

341 Towards uranium catalysts

A. R. Fox, S. C. Bart, K. Meyer & C. C. Cummins

nature
insight

The advent and development of organocatalysis

David W. C. MacMillan

The use of small organic molecules as catalysts has been known for more than a century. But only in the past decade has organocatalysis become a thriving area of general concepts and widely applicable asymmetric reactions. Here I present my opinion on why the field of organocatalysis has blossomed so dramatically over the past decade.

Organocatalysis, or the use of small organic molecules to catalyse organic transformations, is a relatively new and popular field within the domain of chiral molecule (or enantioselective) synthesis. Although chemical transformations that use organic catalysts, or organocatalysts, have been documented sporadically over the past century, it was not until the late 1990s that the field of organocatalysis was 'born', coalescing around a small number of articles that inspired an explosion of research¹. Between 1998 and 2008, at least 1,500 manuscripts describing the use of organocatalysts in more than 130 discrete reaction types were published (Fig. 1), a remarkable number given that there were no reports of such catalysts in the year 1995. It is now widely accepted that organocatalysis is one of the main branches of enantioselective synthesis (the other, previously accepted, branches being enzymatic catalysis and organometallic catalysis; see page 314), and those who are involved in the synthesis of chiral molecules consider organocatalysis to be a fundamental tool in their catalysis toolbox (see page 323).

In this Commentary, I present my opinion on how chemists arrived at this point. It is useful to look back and consider this development, not least because it provides insight into how scientific communities can seemingly bypass an important area of research for many decades and then, at a given moment, embrace the same field with collective (and at times fervent) enthusiasm. I consider that three factors were crucial to the sudden birth and rapid growth of the field of organocatalysis: first, the conceptualization of the field; second, the advantages of organocatalytic research; and, third, the advent of generic modes of catalyst activation, induction and reactivity.

Conceptualization of the field of organocatalysis

New synthetic methods are most likely to be encountered in the fields of biological and organometallic chemistry.

Dieter Seebach (1990)²

Why was organocatalysis so long overlooked as an area of research by the chemical synthesis community? This question is difficult to answer in hindsight, given the now obvious nature of the concepts and catalysts involved. However, one perspective worth considering is that it is impossible to overlook a field that does not yet exist. This might seem cryptic, but it is similar to the idea that researchers cannot work on a problem that has not been identified. Between 1968 and 1997, there were only a few reports of the use of small organic molecules as catalysts for asymmetric reactions (the Hajos–Parrish reaction probably being the most famous), but these chemical studies were viewed more as unique chemical reactions than as integral parts of a larger, interconnected field^{3–5}. In these early publications, there was no emphasis on the potential benefits of using organocatalysts or on the demonstration of

new organocatalytic concepts. Instead, these articles focused only on the individual transformations that had been accomplished. Because general lessons were not extrapolated from these reactions, the underlying catalytic concepts were not applied to multiple reaction types. As a result, chemists did not recognize that the use of small organocatalysts could be an overarching concept. One measure of this fact is that between 1960 and 1998, there were no review articles about the use of organocatalysts. In addition, Dieter Seebach, an organic chemist of great standing, omitted organocatalysis from his vision for synthetic methods in 1990, clearly illustrating that it is not easy to predict the success of a field that, at the time, does not exist.

In the late 1990s, however, things began to change when Yian Shi⁶, Scott Denmark⁷ and Dan Yang⁸, and their co-workers, demonstrated that enantiomerically pure ketones could be used to catalyse the enantioselective epoxidation of simple alkenes. Shortly afterwards, Eric Jacobsen⁹ and Elias J. Corey¹⁰, and their co-workers, described the first examples of hydrogen-bonding catalysis, in an asymmetric Strecker reaction, and Scott Miller and his co-workers¹¹ introduced the concept of minimal peptides for the enantioselective kinetic resolution of alcohols. Although, collectively, this work did not conceptualize organocatalysis as a field of research, it demonstrated for the first time that small organocatalysts could be used to solve important problems in chemical synthesis. It was not until 2000, however, that the field of organocatalysis was effectively launched, by two publications that appeared almost simultaneously: one from Carlos Barbas, Richard Lerner and Benjamin List¹², on enamine catalysis; and the other from my research group¹³, on iminium catalysis.

The work of Barbas, Lerner and List was significant because it showed that the underlying mechanism of the Hajos–Parrish reaction could be extended and applied to transformations that have a broader applicability (specifically, the intermolecular aldol reaction). Moreover, this work showed that small organic molecules (such as proline) could catalyse the same chemical reactions as much larger organic molecules (enzymes) by using similar mechanisms. Meanwhile, the report of iminium catalysis conceptualized 'organocatalysis' in three important ways: by delineating how organocatalysts could provide economic, environmental and scientific benefits; by describing a general activation strategy for organocatalysis that could be applied to a broad range of reaction classes (rather than simply allowing a single transformation); and by introducing the term organocatalysis to the chemical literature. You might ask, What's in a name? But consider the success of the terms nanotechnology and diversity-oriented synthesis at globally shifting the visibility and perception of areas of research. The term organocatalysis provided a strong identity and helped to unify a fledgling field, as well as attracting the attention of the broader chemical synthesis community. At the last count, the word organocatalysis, or one of its derivatives, has been used

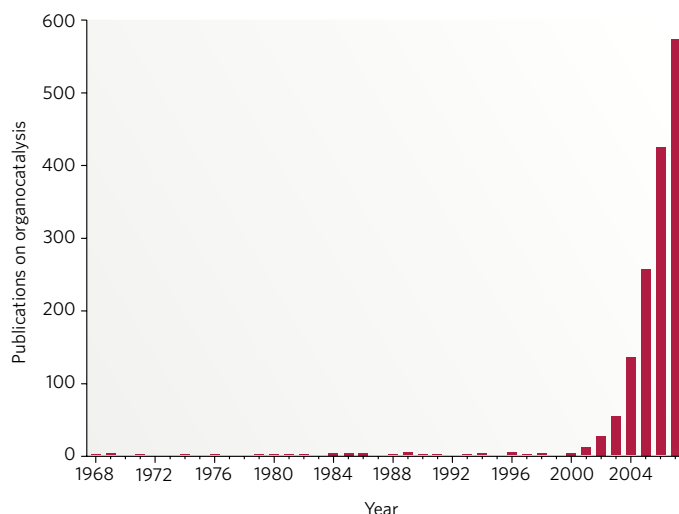


Figure 1 | An explosion of interest. The number of publications on the topic of organocatalysis has recently increased markedly. Data were obtained by a search of the ISI Web of Knowledge in May 2008 for the keyword organocatalysis and its derivatives, as well as the keywords phase transfer, enamine and Hajos (to locate articles on this topic before the term organocatalysis was coined). This search is unlikely to have found all publications on organocatalysis, and a conservative estimate is that more than 2,000 manuscripts have been published on this topic so far.

in more than 600 article titles and is the subject of more than 40,000 web pages (based on a survey of the ISI Web of Knowledge and of web pages indexed by Google in May 2008).

Advantages of organocatalytic research

Once the field of organocatalysis had been publicly defined, it grew quickly. Such rapid adoption is possible only when a field offers real advantages to the researchers who mine the area and to those who exploit the resultant technologies. In this case, the chemical synthesis community quickly recognized the fundamental advantages of organocatalysis, namely the ease and low cost of carrying out such reactions in the laboratory and the potential for new lines of academic thought and investigation. Before 1998, the state of the art in asymmetric catalysis involved metal-based chiral catalysts almost exclusively, and these catalysts allowed a wealth of oxidations, reductions, σ -bond insertions, π -bond activations and Lewis-acid-catalysed reactions. Although the impact of metal-based catalysts on chemical synthesis cannot be understated, some (but certainly not all) organometallic systems can be expensive, toxic and/or sensitive to air and moisture.

The advent of organocatalysis brought the prospect of a complementary mode of catalysis, with the potential for savings in cost, time and energy, an easier experimental procedure, and reductions in chemical waste (Fig. 2). These benefits arise from three factors. First, organic molecules are generally insensitive to oxygen and moisture in the atmosphere, so there is no need for special reaction vessels, storage containers and experimental techniques, or for ultra-dry reagents and solvents. Second, a wide variety of organic reagents — such as amino acids, carbohydrates and hydroxy acids — are naturally available from biological sources as single enantiomers. Simple organocatalysts are therefore usually cheap to prepare and readily accessible in a range of quantities, suitable for small-scale reactions to industrial-scale reactions. Third, small organic molecules are typically non-toxic and environmentally friendly, increasing the safety of catalysis in both biological research and chemical research across all research settings, including industry and academic institutions.

The combination of these factors substantially lowered the entry costs for researchers who were interested in developing enantioselective catalysts. With no need for gloveboxes, inert gases, ultra-dry solvents or even high levels of experimental expertise, it is not surprising that the field quickly became flooded with research groups from around the globe.

Indeed, the increase in competition helped to accelerate the pace of innovation and discovery, albeit perhaps at the expense of 'elbow room' for those in this popular field.

Some critics suggest that low turnover numbers might limit the potential uses of organocatalysis for industrial applications, but this view is simplistic and dogmatic. For any large-scale catalytic process, the most salient considerations are cost and safety. Because organocatalysts are often cheaper than metal-based catalysts, organocatalysts can be used in larger quantities than metal-based ones for the same price. Moreover, it is widely recognized in manufacturing that the removal of toxic catalyst-related impurities from the waste stream can often have a larger financial impact than the turnover number of the catalyst. Organocatalysts are typically less toxic than metal-based catalysts, can be tolerated to a large extent in waste streams and are more easily removed from waste streams, again mitigating the cost of high catalyst loadings.

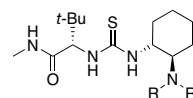
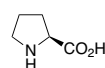
With respect to industrial applications, organocatalytic reactions are of great use to medicinal chemists. These researchers are the largest body of industrial synthetic chemists in the world, and their technological needs often differ from those of chemists in manufacturing or process development. Medicinal chemists need to find rapid, broadly applicable ways of constructing new candidate drugs for testing, so the most important considerations for a catalyst are its generality, convenience and robustness. Organocatalysts meet all of these operational requirements. Indeed, organocatalysts have already been taken up by some medicinal chemists in pursuit of therapeutic agents enriched for a particular enantiomer¹⁴.

Advent of generic modes of activation and induction

Perhaps most crucial to the success of organocatalysis in the past decade has been the invention or identification of generic modes of catalyst activation, induction and reactivity. A generic activation mode describes a reactive species that can participate in many reaction types with consistently high enantioselectivity (as opposed to one or two unique transformations). Such reactive species arise from the interaction of a single chiral catalyst with a basic functional group (such as a ketone, aldehyde, alkene or imine) in a highly organized and predictable manner.

The value of generic activation modes is that, after they have been established, it is relatively straightforward to use them as a platform for designing new enantioselective reactions. Indeed, most of the 130 organocatalytic reactions that have been reported since 1998 are founded directly on only five or six activation modes. The small number of activation modes in organocatalysis (and in catalysis in general) is not surprising — when devising a new enantioselective reaction, it is far easier to make use of a known activation mode than to invent a new one (together with a new catalyst). Hence, the number of enantioselective catalytic reactions will always be much larger than the number of activation modes that underpin them.

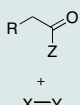
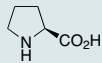
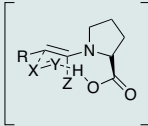

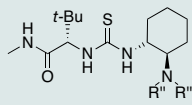
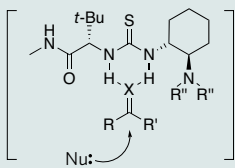

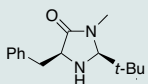
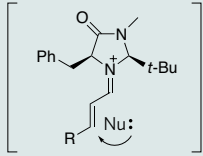

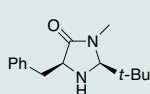
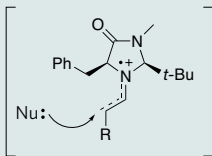

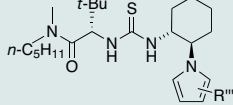
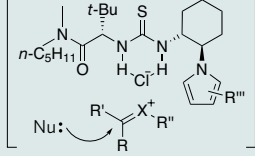
In many ways, this is beneficial to chemists, because it leads to the development of catalysts or catalyst families that are useful for a wide range of asymmetric reactions. At the same time, it is clear that discovering new activation modes is important for all types of catalysis. In the field of organocatalysis, iminium catalysis, reported in 2000, was the first



- Stable in air and water
- Available from biological materials
- Inexpensive and easy to prepare
- Simple to use
- Both enantiomers available
- Non-toxic

Figure 2 | Typical organocatalysts and the advantages of using organocatalysts. L-Proline (top) is often used as an organocatalyst for many reactions, notably for functionalizing carbonyl-containing molecules at the α -carbon. The mechanism involves the formation of a reactive enamine intermediate. Thioureas (bottom) catalyse several reactions by forming hydrogen bonds with the substrate. The typical advantages of using organocatalysts are listed. R, alkyl; *t*-Bu, *tert*-butyl.

Table 1 | Generic modes of activation commonly used in organocatalysis

Substrate	Catalyst	Activation mode	Number of new reactions	Examples of new reaction variants	References
Enamine catalysis					
 $R = \text{any organic chain or ring system}$ $X = C, N, O, S$ $Y = \text{generic organic atom}$ $Z = \text{alkyl, H}$		HOMO activation 	25	<ul style="list-style-type: none"> Aldehyde-aldehyde cross aldol coupling Intramolecular α-alkylation Mannich reaction Michael reaction α-Amination α-Oxygenation α-Halogenation α-Sulphenylation 	First application: 1971, ref. 3 First use as a generic mode of activation: 2000, ref. 13
Hydrogen-bonding catalysis					
 $X = O, NR$ $R, R', R'' = \text{alkyl, aryl}$		LUMO activation 	30	<ul style="list-style-type: none"> Strecker reaction Mannich reaction Ketone cyanosilylation Biginelli reaction Pictet-Spengler reaction Reductive amination 	First application: 1996–1999, refs 9–11 First use as a generic mode of activation: 2002, ref. 17
Iminium catalysis					
 $R = \text{alkyl, aryl}$		LUMO activation 	50	<ul style="list-style-type: none"> Conjugate Friedel-Crafts reaction Ketone Diels-Alder reaction exo-Selective Diels-Alder reaction Mukaiyama-Michael reaction Conjugate hydride reduction Conjugate amination Conjugate oxygenation Conjugate sulphenylation Cyclopropanation Epoxidation, aziridination 	First application and first use as a generic mode of activation: 2000, ref. 14
SOMO catalysis					
 $R = \text{alkyl, aryl}$		SOMO activation 	4	<ul style="list-style-type: none"> α-Allylation α-Enolation α-Vinylolation α-Heteroarylation 	First application and first use as a generic mode of activation: 2007, ref. 25
Counterion catalysis					
 $X = O, NR$ $R, R', R'', R''' = \text{alkyl, aryl}$		LUMO activation 	2	<ul style="list-style-type: none"> Acyl-Pictet-Spengler reaction Oxocarbenium addition reaction 	First application and first use as a generic mode of activation: 2007, ref. 28

In HOMO activation, the energy of the highest-occupied molecular orbital (HOMO) is increased. In LUMO activation, the energy of the lowest-unoccupied molecular orbital (LUMO) is decreased. In SOMO activation, an electron is located in a singly occupied molecular orbital (SOMO) to generate a highly reactive species that can participate in many reaction types. Nu, nucleophile; Ph, phenyl.

activation mode identified as being suitable for more than one reaction type, a feature that was central to its design¹³. Enamine catalysis was also demonstrated to be a generic activation mode by its application to the Mannich reaction in 2000 (ref. 15), and Jacobsen and Anna Wenzel were the first to demonstrate that hydrogen-bonding catalysis had multiple-reaction utility, in 2002 (ref. 16). The activation modes discussed in this section and described in Table 1 are representative of modes that have been invented or identified in the past decade.

Enamine catalysis

In 1971, there were two independent reports — one by Zoltan Hajos and David Parrish³, and another by Rudolf Weichert, Gerhard Sauer and Ulrich Eder⁴ — of an enantioselective intramolecular aldol reaction that was catalysed by proline in the synthesis of the Wieland–Miescher ketone. This extraordinary result was well received by the community, but the underlying activation mode was not exploited for other reactions until more than 30 years later. It was in 2000, with the ingenious work of

Barbas, Lerner and List — who used enamine catalysis to functionalize carbonyl-containing compounds at the α -carbon — that the broad applicability of this mode of activation became evident¹². Since then, a tremendous amount of research has been directed towards identifying new types of chiral enamine catalysts.

Mechanistically, enamine catalysis might be better described as bifunctional catalysis because the amine-containing catalyst (proline in Table 1) typically interacts with a ketone substrate to form an enamine intermediate but simultaneously engages with an electrophilic reaction partner through either hydrogen bonding or electrostatic attraction. This mode of activation has now been used in a wide range of enantioselective carbonyl α -functionalization processes¹⁷.

Hydrogen-bonding catalysis

In the early 1980s, researchers uncovered several catalytic asymmetric processes that suggested that the activation of a substrate and the organization of the transition state could occur through well-defined hydrogen-bonding interactions^{18–20}. These reports were widely appreciated; however, they were considered to be exceptions to the generally held idea that hydrogen bonding was insufficiently activating or directional for use in asymmetric catalysis. This reasoning was disproved by reports published in 1998 and 1999, when Jacobsen⁹ and Corey¹⁰ independently reported an asymmetric variant of the Strecker reaction that used well-defined hydrogen-bonding organocatalysts that activate imine electrophiles. Moreover, four years later, Jacobsen showed that these thiourea catalysts could be used for other synthetic reactions¹⁶, launching the generic use of enantioselective hydrogen-bonding catalysis. This powerful activation mode has become the foundation of a large and dynamic area of research. More than 30 new asymmetric reactions have been developed on the basis of this principle²¹.

Iminium catalysis

Iminium catalysis was the first organocatalytic activation mode to be designed (rather than discovered) and introduced as a general strategy for asymmetric organic synthesis. It is based on the capacity of chiral amines to function as enantioselective catalysts for several transformations that traditionally use Lewis acid catalysts. The concept was founded on the mechanistic hypothesis that the reversible formation of iminium ions from α,β -unsaturated aldehydes and chiral amines might emulate the equilibrium dynamics and π -orbital electronics that are inherent to Lewis acid catalysis (that is, lowest-unoccupied molecular orbital (LUMO)-lowering activation). With its tailor-made family of imidazolidinone catalysts, iminium catalysis is now used in more than 50 highly enantioselective protocols²², many of which have been developed by my research group and that of Karl Anker Jørgensen.

SOMO catalysis

Introduced by my research group in 2006, SOMO catalysis is based on the idea that one-electron oxidation of an electron-rich enamine selectively generates a reactive radical cation with three π -electrons. The electrophilicity of the singly occupied molecular orbital (SOMO) of this intermediate allows it to react readily with a variety of weakly nucleophilic carbon-based 'SOMOPhiles' at the α -carbon of the parent enamine, resulting in formal alkylation products²³. Applying this principle in a catalytic system that uses chiral secondary amines and a suitable one-electron oxidant has been highly successful, opening up new avenues for catalysing the asymmetric α -functionalization of carbonyl-containing compounds. Although SOMO catalysis is one of the most recently discovered activation modes, it has already yielded a series of enantioselective transformations that complement those produced by enamine catalysis^{24–26}.

Counterion catalysis

Jacobsen recently developed a conceptually novel form of organocatalytic activation that directs highly enantioselective additions into transiently generated *N*-acyl-iminium ions and oxocarbenium ions^{27,28}. In this system, chiral thiourea catalysts, which are known to form strong complexes with halide ions, electrostatically bind to, and ionize, the weak

carbon–chlorine bonds of chloroamides and chloroacetals to generate a transient ion pair. The resultant anionic catalyst–chloride complex functions as a chiral counterion, biasing the approach of nucleophiles to a single face of one enantiomer of the transient α -heteroatom-stabilized cationic species. The initial results are remarkable in that forces acting through space, rather than through bonds, are sufficient to transfer stereochemical information from the catalyst to the substrate with high fidelity. Although this mode of activation is still in its infancy, it has the potential to solve many persistent and outstanding problems in asymmetric catalysis.

The future

It is always a risk to predict what the future holds for any field, but several aspects of organocatalysis will undoubtedly attract researchers' attention. Tremendous efforts will continue to be directed towards the discovery and design of catalysts with better efficiency, new reactivities and greater turnover numbers. In addition, although SOMO catalysis and counterion catalysis have been developed only recently, it is clear that at least ten new asymmetric transformations will arise from these modes of activation. Moreover, given the substantial impact of new generic modes of activation, it is probable that many young and ambitious chemists will focus their research efforts on attaining this ideal. More specific challenges that are likely to be addressed include the organocatalytic activation of carbon–hydrogen bonds and the catalysis of photoredox reactions. There will also probably be further advances in the area of organocatalytic cascade reactions, in which activation modes are combined to provide powerful transformations that generate complex molecules^{29,30}. Although cascade catalysis is rarely used at present, it is likely to become more prevalent for the synthesis of natural products and medicinal agents. Furthermore, it is probable that computational models will increasingly be used to aid in not only the understanding of organocatalytic processes but also the design of new organocatalysts (see page 309).

But perhaps the most crucial area of research in the future will be the identification of important transformations and new reactivities that are not available using other branches of catalysis. Given the huge growth and impact of organocatalysis in the past decade, it will certainly be exciting to observe the development of the field over the next decade.

David W. C. MacMillan is at the Merck Center for Catalysis at Princeton University, 116 Frick Laboratory, Princeton University, Princeton, New Jersey 08540, USA.

1. Berkessel, A. & Groeger, H. *Asymmetric Organocatalysis: from Biomimetic Concepts to Applications in Asymmetric Synthesis* (Wiley-VCH, 2005).
2. Seebach, D. Organic synthesis — where now? *Angew. Chem. Int. Edn Engl.* **29**, 1320–1367 (1990).
3. Hajos, Z. G. & Parrish, D. R. Asymmetric synthesis of optically active polycyclic organic compounds. German patent DE 2102623 (1971).
4. Eder, U., Sauer, G. R. & Wiechert, R. Optically active 1,5-indanone and 1,6-naphthalenedione derivatives. German patent DE 2014757 (1971).
5. Hajos, Z. G. & Parrish, D. R. Asymmetric synthesis of bicyclic intermediates of natural product chemistry. *J. Org. Chem.* **39**, 1615–1621 (1974).
6. Tu, Y., Wang, Z. & Shi, Y. An efficient asymmetric epoxidation for *trans*-olefins mediated by a fructose derived ketone. *J. Am. Chem. Soc.* **118**, 9806–9807 (1996).
7. Denmark, S. E., Wu, Z., Crudden, C. & Matsushashi, H. Catalytic epoxidation of alkenes with oxone. 2. Fluoro ketones. *J. Org. Chem.* **62**, 8288–8289 (1997).
8. Yang, D. et al. A C_2 symmetric chiral ketone for catalytic asymmetric epoxidation of unfunctionalized olefins. *J. Am. Chem. Soc.* **118**, 491–492 (1996).
9. Sigman, M. & Jacobsen, E. N. Schiff base catalysts for the asymmetric Strecker reaction identified and optimized from parallel synthetic libraries. *J. Am. Chem. Soc.* **120**, 4901–4902 (1998).
10. Corey, E. J. & Grogan, M. J. Enantioselective synthesis of α -amino nitriles from *N*-benzhydryl imines and HCN with a chiral bicyclic guanidine as catalyst. *Org. Lett.* **1**, 157–160 (1999).
11. Miller, S. J. et al. Kinetic resolution of alcohols catalyzed by tripeptides containing the *N*-alkylimidazole substructure. *J. Am. Chem. Soc.* **120**, 1629–1630 (1998).
12. List, B., Lerner, R. A. & Barbas, C. F. III. Proline-catalyzed direct asymmetric aldol reactions. *J. Am. Chem. Soc.* **122**, 2395–2396 (2000).
13. Ahrendt, K. A., Borths, C. J. & MacMillan, D. W. C. New strategies for organic catalysis: the first highly enantioselective organocatalytic Diels–Alder reaction. *J. Am. Chem. Soc.* **122**, 4243–4244 (2000).
14. King, H. D. et al. Enantioselective synthesis of a highly potent selective serotonin reuptake inhibitor. An application of imidazolidinone catalysis to the alkylation of indoles with an α,β -disubstituted α,β -unsaturated aldehyde. *Org. Lett.* **7**, 3437–3440 (2005).
15. List, B. The direct catalytic asymmetric three-component Mannich reaction. *J. Am. Chem. Soc.* **122**, 9336–9337 (2000).

16. Wenzel, A. G. & Jacobsen, E. N. Asymmetric catalytic Mannich reactions catalyzed by urea derivatives: enantioselective synthesis of β -aryl- β -amino acids. *J. Am. Chem. Soc.* **124**, 12964–12965 (2002).
17. Mukherjee, S., Yang, J. W., Hoffmann, S. & List, B. Asymmetric enamine catalysis. *Chem. Rev.* **107**, 5471–5569 (2007).
18. Hiemstra, H. & Wynberg, H. Addition of aromatic thiols to conjugated cycloalkenones, catalyzed by chiral β -hydroxy amines. A mechanistic study on homogeneous catalytic asymmetric synthesis. *J. Am. Chem. Soc.* **103**, 417–430 (1981).
19. Oku, J. I. & Inoue, S. Asymmetric cyanohydrin synthesis catalyzed by a synthetic cyclic dipeptide. *J. Chem. Soc. Chem. Commun.* 229–230 (1981).
20. Dolling, U. H., Davis, P. & Grabowski, E. J. J. Efficient catalytic asymmetric alkylations. 1. Enantioselective synthesis of (+)-indacrinone via chiral phase-transfer catalysis. *J. Am. Chem. Soc.* **106**, 446–447 (1984).
21. Doyle, A. G. & Jacobsen, E. N. Small-molecule H-bond donors in asymmetric catalysis. *Chem. Rev.* **107**, 5713–5743 (2007).
22. Lelais, G. & MacMillan, D. W. C. Modern strategies in organic catalysis: the advent and development of iminium activation. *Aldrichim. Acta* **39**, 79–87 (2006).
23. Narasaka, K., Okauchi, T., Tanaka, T. & Murakami, M. Generation of cation radicals from enamines and their reactions with olefins. *Chem. Lett.* **92**, 2099–2102 (1992).
24. Beeson, T. D., Mastracchio, A., Hong, J., Ashton, K. & MacMillan, D. W. C. Enantioselective organocatalysis using SOMO activation. *Science*, **316**, 582–585 (2007).
25. Jang, H., Hong, J. & MacMillan, D. W. C. Enantioselective organocatalytic singly occupied molecular orbital activation: the enantioselective α -enolization of aldehydes. *J. Am. Chem. Soc.* **129**, 7004–7005 (2007).
26. Kim, H. & MacMillan, D. W. C. Enantioselective organo-SOMO catalysis: the α -vinylolation of aldehydes. *J. Am. Chem. Soc.* **130**, 398–399 (2008).
27. Raheem, I., Thiara, P. S., Peterson, E. A. & Jacobsen, E. N. Enantioselective Pictet–Spengler-type cyclizations of hydroxylactams: H-bond donor catalysis by anion binding. *J. Am. Chem. Soc.* **129**, 13404–13405 (2007).
28. Reisman, S. E., Doyle, A. G. & Jacobsen, E. N. Enantioselective thiourea-catalyzed additions to oxocarbenium ions. *J. Am. Chem. Soc.* **130**, 7198–7199 (2008).
29. Huang, Y., Walji, A. M., Larsen, C. H. & MacMillan, D. W. C. Enantioselective organo-cascade catalysis. *J. Am. Chem. Soc.* **127**, 15051–15053 (2005).
30. Enders, D., Grondal, C. & Huttel, M. R. M. Asymmetric organocatalytic domino reactions. *Angew. Chem. Int. Edn Engl.* **46**, 1570–1581 (2007).

Acknowledgements I thank the National Institute of General Medical Sciences (grant number R01 GM078201-01-01) for financial support, and Merck Research Laboratories and Bristol-Myers Squibb for gifts. I also thank H. Kim, R. Knowles, D. Carerra and J. Van Humbeck for their help in the preparation of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to the author (dmacmill@princeton.edu).

Computational prediction of small-molecule catalysts

K. N. Houk & Paul Ha-Yeon Cheong

Most organic and organometallic catalysts have been discovered through serendipity or trial and error, rather than by rational design. Computational methods, however, are rapidly becoming a versatile tool for understanding and predicting the roles of such catalysts in asymmetric reactions. Such methods should now be regarded as a first line of attack in the design of catalysts.

Just 40 years ago, computational methods for simulating chemical reactions used in synthesis did not exist. Since then, the discovery of several computational principles and algorithms — together with the development of fast computers — has resulted in enormous leaps in the accuracy and speed of computational methods, and it is now feasible to model many synthetic reactions in this way. It was the release of Gaussian 70 (a general-purpose, *ab initio* computer program that has developed into today's Gaussian 03 software)¹ and the advent of minicomputers, both in the 1970s, that brought the tools of quantum mechanics to organic chemists. The subsequent development of quantum-mechanical techniques such as density-functional-theory calculations² and high-accuracy *ab initio* computational methods proceeded at an ever-accelerating pace³. In addition, methods were devised to reduce the lengthy times required to model large systems (those containing thousands of atoms) or smaller systems with high accuracy, and these methods continue to be improved³.

The cost of computation has also decreased enormously since the 1960s. Moore's law — which states that the number of transistors on an integrated circuit doubles every two years — reflects the incredible rate at which computing power is increasing. But a more impressive figure involves the economics of computing: the cost of 1 gigaflops of computing power in the 1960s was about US\$7.9 trillion in today's currency, whereas the same computing power is available now for less than 20 cents (ref. 4 and see <http://news.bbc.co.uk/2/hi/technology/4554025.stm>).

Today, there are computational methods for modelling molecules and reactions that can obtain results at almost any accuracy desired, as long as sufficient computer resources are available (Fig. 1). In practice, the energies of small molecular systems can be computed at greater accuracies (± 1 kilocalorie per mol) than those achievable in most experiments. Larger systems are generally modelled at a lower accuracy so that results can be obtained within a reasonable time. The errors in absolute energies (typically ± 5 kilocalories per mol) might therefore be larger than desired, but a much greater accuracy can be achieved when comparing several similar structures, for example the transition states of diastereomeric molecules.

Computational methods are proving to be especially useful for modelling catalytic reactions. Such theoretical approaches readily provide insights into the mechanisms of catalytic processes and their selectivities for various products. Indeed, a theoretical approach is sometimes faster than an experimental one and is often the only way to attain the level of detail that chemists seek. Small chiral catalysts — either organic catalysts (organocatalysts; see page 304) or transition-metal complexes (organometallic catalysts; see page 314) — are ideal for computational modelling, because their size makes accurate simulations feasible, allowing detailed results to be generated and used to design new catalysts.

Predicting how effective such catalysts will be in a specific reaction is still a daunting task, but several notable successes make it clear that computational design is becoming an increasingly powerful tool. We argue here that computational design is ready to take its place as an essential component of catalyst design.

The road to catalyst design

Most computational chemistry research involves the explanation of phenomena that are not well understood. Current quantum-mechanical methods have been successfully used for such rationalization of chemical phenomena. But, as quantum-mechanical methods have developed, it has become increasingly possible to make predictions (rather than rationalizations), and several predictions have now been confirmed experimentally. The first such predictions were for non-catalytic reactions involving small molecules. An example is the prediction and subsequent experimental verification of contra-steric stereochemistry in the products of the electrocyclic ring-opening of 3-formylcyclobutene⁵ (Fig. 2), a non-intuitive result that would not have been predicted by researchers on the basis of knowledge about reactions of the analogous molecule 3-methylcyclobutene. Subsequent to this prediction, Masahiro Murakami and colleagues predicted that, in cyclobutene ring-opening reactions, silyl groups and stannyl groups would also anomalously rotate inwards, and they then verified these predictions experimentally^{6,7}. Similarly, unexpected outcomes have been predicted computationally and verified experimentally for reactions commonly used in synthesis; for example, the stereoselectivity of hydride additions to cycloheptanones has been computed and confirmed in this way⁸.

The first organocatalytic reaction to be studied computationally was the Hajos–Parrish reaction^{9–11} (an asymmetric, intramolecular aldol reaction that is catalysed by L-proline), and this work was pivotal in showing the power of computational methods for studying and predicting catalytic reactions. Four mechanisms had been proposed for the Hajos–Parrish reaction. Of these, computations revealed a preference for the *s-trans* transition state (which has a vinyl group and a carboxylic acid in a *trans* arrangement across an enamine single bond), which leads to the experimentally observed product¹² (Fig. 3). The computed energy of 3.5 kilocalories per mol favouring the formation of the major stereoisomer corresponded well to the experimentally observed stereoselectivity of 95% enantiomeric excess, and could be explained by analysing the structural features that stabilize the *s-trans* transition structure rather than the *s-cis* form. Building on this precedent, computational studies of aldol reactions catalysed by proline derivatives have now been published¹³, including some predictions of new catalysts — although these have yet to be confirmed experimentally^{14,15}.

Designer organocatalysts realized

The first and most successful examples of the computational design of a stereoselective organocatalyst that was later experimentally validated came from studies of highly selective *anti* Mannich reactions¹⁶. Proline-catalysed Mannich reactions are selective for products in the *syn* orientation, but Carlos Barbas and his colleagues discovered that by using pipelicolic acid (a proline analogue with a six-membered ring) as the catalyst, a mixture of *syn* and *anti* products resulted. The challenge then was whether further modifications could be made to the catalyst to result in reactions that were highly selective for the *anti* orientation.

In collaboration with Barbas's research group, we modelled the transition states of these reactions and found that selectivity for the *anti* orientation arises when the *s-cis*-enamine intermediate (formed from a reaction between the catalyst and the aldehyde reactant) attacks the imine starting material, rather than the alternative possibility in which the *s-trans*-enamine attacks¹⁷. On this basis, several potential catalysts that would favour the *s-cis* transition state were proposed, and computations were carried out to predict which of these molecules would be highly selective for the product in the *anti* orientation. The molecule shown in Fig. 4a was predicted to result in the *anti* product selectively (with 5% *syn* products and 95% *anti* products, and 98% enantiomeric excess). Subsequent experiments with this molecule by Barbas's research group showed remarkable agreement with these predictions (6% *syn* and 94% *anti*, and 99% enantiomeric excess)¹⁶. Such *in silico* screening can be much quicker than making each potential catalyst and checking it experimentally, so these studies show how computational methods could be used to save time when designing catalysts.

Other types of organocatalysis have also proved amenable to computational predictions. Thioureas, for example, are effective catalysts for activating various substrates containing a carbonyl group¹⁸. But they are not as effective at activating epoxides. To tackle this problem, Stephen

Connon, Isabel Rozas and their colleagues used computational methods to model the structures of analogues of thioureas and ureas in complex with epoxides¹⁹. On the basis of their results, they predicted that *N*-tosyl ureas would form stronger complexes with epoxides than would thioureas or ureas and therefore could catalyse epoxide ring-opening reactions. Subsequent experiments showed that this was the case — *N*-tosyl ureas successfully catalyse the addition of indoles to styrene oxides (Fig. 4b), whereas the previously known bis-(aryl)thiourea complexes are essentially inert.

These are examples of the state of the art in the computational design of catalysts. The greatest successes have involved modifying molecules that show some catalytic activity, through using the knowledge gleaned from computational models to design a second generation of catalysts. Given a well-understood candidate catalyst, chemists can use their intuition to suggest modified compounds that might work better or a diverse library of structurally related analogues to explore. Computations can then be used to determine which of these should be the best catalyst. This is akin to computational approaches to drug design, in which the activity of a compound is successfully optimized by starting with a 'lead' structure.

Designing organometallic reactions

Computational approaches are most accurate for predicting the catalytic activities of analogues of existing catalysts. But, in theory, these approaches should be able to make predictions about any class of compound, not just analogues of those used as a 'training set' for the computational model. Marisa Kozlowski and her colleagues have developed several computational tools to achieve this goal. They have created computational models for catalytic activity based on either the transition state for a particular reaction²⁰ (as calculated by quantum-mechanical methods) or the structures of the catalysts in their ground states²¹ (which are much easier to determine than transition states). Quantitative

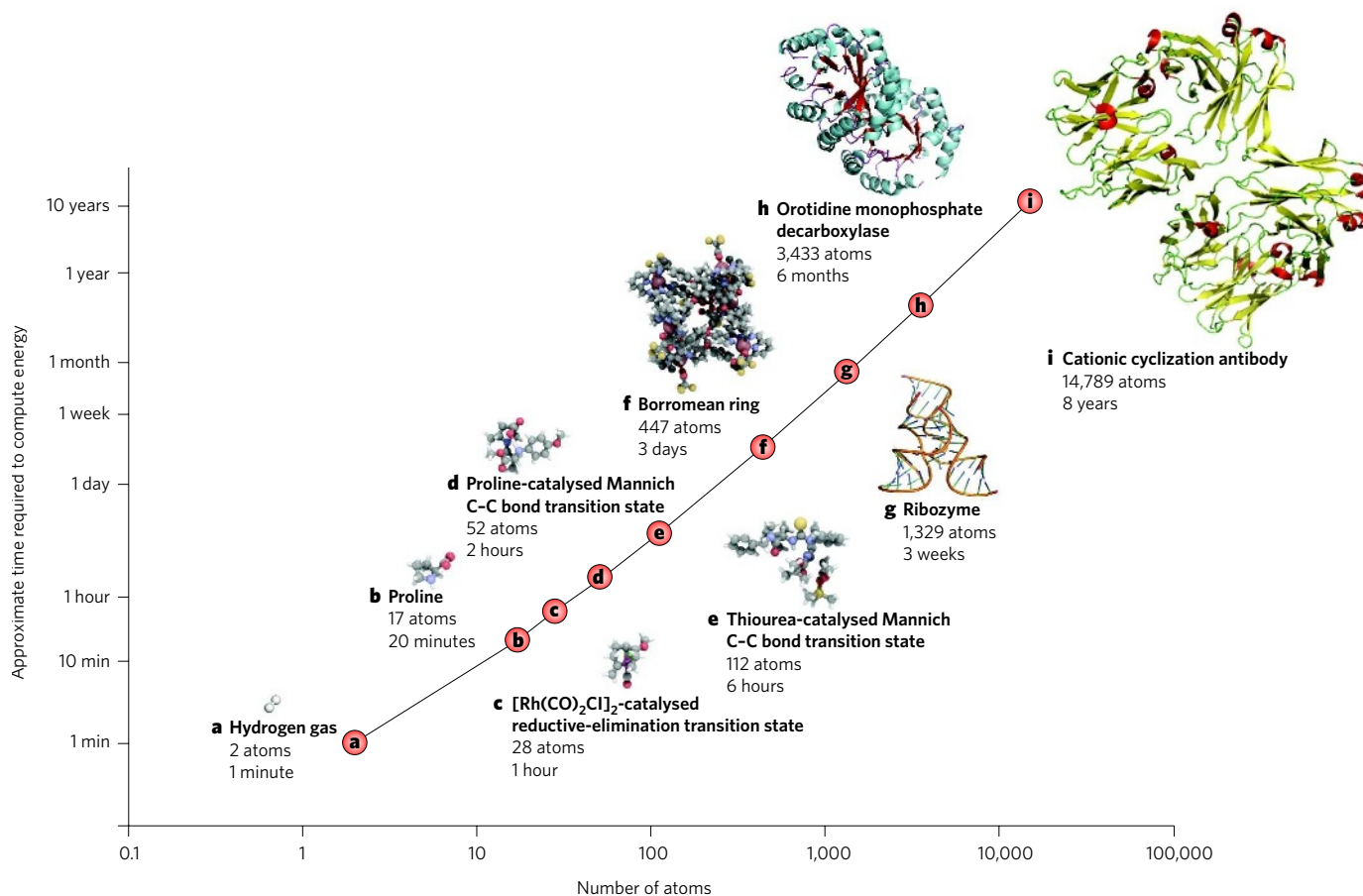


Figure 1 | Time required to compute molecular structures. Shown is the approximate time taken to compute the structure of molecules of various sizes by using density functional theory and double-zeta basis sets on a modern desktop computer with appropriate memory and storage. The time required (shown on a log scale) increases exponentially with the size of the molecule. Structures were generated using PyMOL³⁵, QuteMol³⁶ or CYLview³⁷.

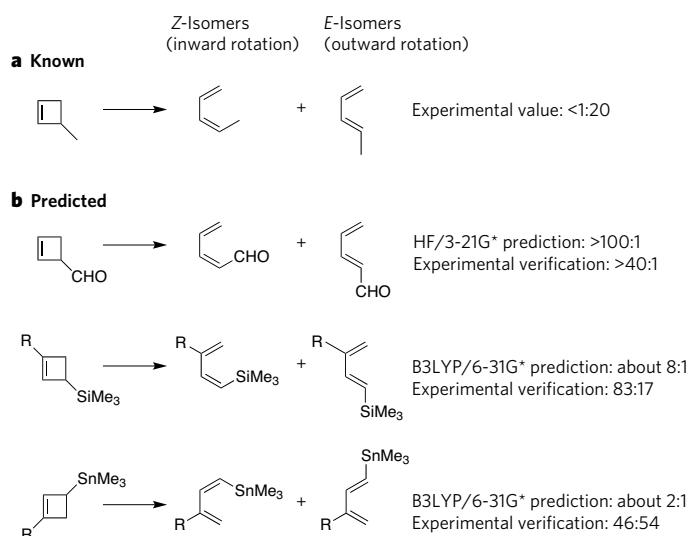


Figure 2 | Successful computational predictions of non-catalytic reactions.

a, In cyclobutene ring-opening reactions, 3-methylcyclobutene forms mainly the *E*-isomer of 1-methylbutadiene (as well as a small amount of the *Z*-isomer); the experimentally determined ratio is <1 *Z*-isomer to 20 *E*-isomers. **b**, Using this knowledge as a starting point, computations correctly predicted the stereochemical outcomes of three analogous reactions^{5–7}, as verified by subsequent experiments, even though the products were mainly *Z*-isomers rather than *E*-isomers. In these reactions, the electron-withdrawing substituents favour inward rotation, whereas, in the original reaction, the electron-donating substituent favours outward rotation. HF/3-21G* is an *ab initio* computational method, and B3LYP/6-31G* is a density-functional-theory computational method. Me, methyl; R, CMe₂Ph.

structure–selectivity relationships for catalysts can thus be developed. This method was used to predict new and effective β -aminoalcohols, ligands that help to catalyse enantioselective additions of diethylzinc to benzaldehyde.

Computational methods not only can help to design catalysts for known reactions but also can predict new catalytic reactions. Paul Wender and his colleagues have developed a series of reactions that are

catalysed by the rhodium dimer [Rh(CO)₂Cl]₂, including several [5+2]-cycloaddition reactions of vinylcyclopropanes with alkynes or allenes²². Intriguingly, the analogous reactions with alkenes do not work. From this collaboration²² with Wender and one of us (K.N.H.), Zhi-Xiang Yu knew that the problem lay in the final step of the reaction mechanism (a reductive-elimination reaction). He speculated that an alternative alkene reaction — a [5+2+1]-multicomponent cycloaddition to form cyclooctenones — might occur if carbon monoxide was added to the reaction mixture. Yu computed the energies of the proposed new reaction and found that the [5+2+1]-cycloaddition was feasible²³. He and his co-workers then confirmed this prediction experimentally (Fig. 4c), showing that cyclooctenones form in the presence of carbon monoxide.

Speed versus accuracy

From the studies discussed here, it is clear that computational chemistry has proved its worth as a predictive tool for the development of catalysts. Methods for rapidly screening virtual libraries of catalysts have been developed²⁴, and leads from completed experimental screens can be modified *in silico*. Thus, only the most effective compounds need to be pursued in follow-up experiments. Nevertheless, highly accurate methods for modelling large molecular systems are lacking, and certain compromises must, inevitably, be made in terms of the accuracy of theoretical models. The development of techniques that allow reasonably accurate predictions in a short time is therefore a crucial strand of current research.

One example of such work has been reported by Ghislain Deslongchamps and his colleagues. They have developed a rapid computational method to screen transition structures of catalysed processes and have applied this method to various reactions, such as the TADDOL-catalysed hetero-Diels–Alder reaction²⁵. This useful method is based on molecular mechanics and assumes that the geometry of the substrate in the uncatalysed process is the same as in the catalysed one. However, this assumption might not always be accurate, so care must be taken in interpreting the results. Such methods sometimes predict transition states with relative energies that differ from those obtained using fully quantum-mechanical methods, owing to the approximate nature of the force-field method²⁶.

Experimental studies are usually carried out in solution, and the calculation of solvation energies is another factor that reduces the accuracy of

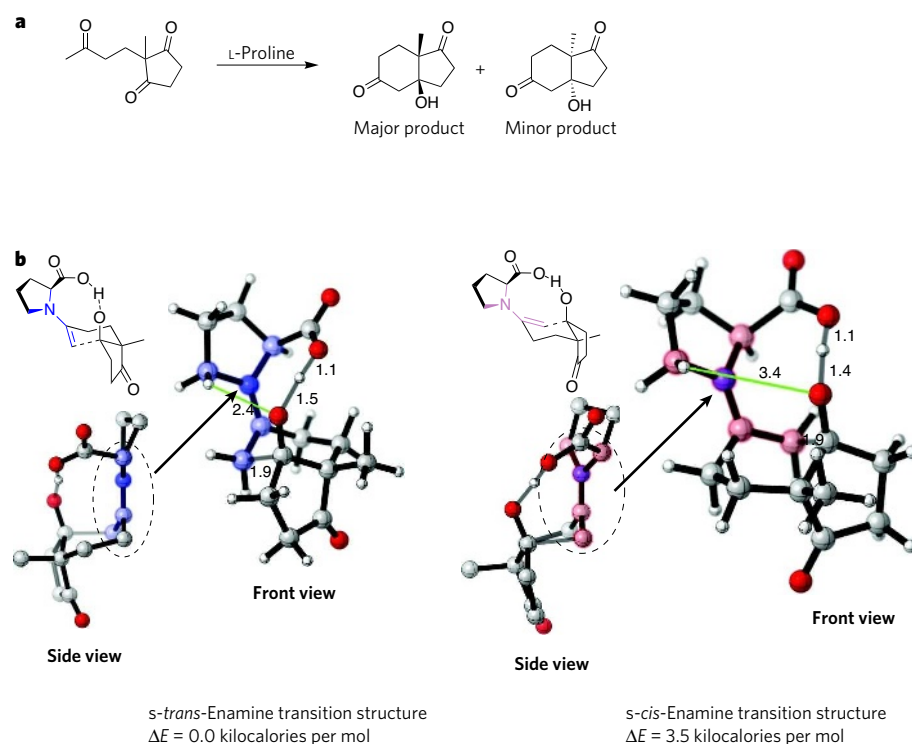


Figure 3 | Transition structures of the Hajos–Parrish reaction.

The Hajos–Parrish reaction (**a**) was one of the first reported organocatalytic processes. Four possible transition states for this reaction were modelled¹² using the hybrid density-functional-theory method B3LYP and the basis set 6-31G*. The computations revealed that the most favoured transition state involves the *s-trans*-enamine intermediate (**b**, left), which incorporates a shorter, and therefore more strongly stabilizing, electrostatic interaction (green) between a hydrogen atom in the pyrrolidine ring and the oxygen atom of the developing alkoxide than does the *s-cis*-enamine intermediate (**b**, right). The side view of the *s-trans* enamine shows that this intermediate is a stable, undistorted (planar) developing iminium ion, whereas the *s-cis*-enamine intermediate is an unstable, distorted (non-planar) developing iminium ion. The favoured transition state leads to the formation of the major product that is observed experimentally, whereas the other transition state leads to the minor product. The main structural differences between the two intermediates are shaded in pale blue (*s-trans* form) and pale pink (*s-cis* form). Carbon is shown in grey, hydrogen in white, oxygen in red and nitrogen in blue. Distances between atoms are indicated in angstroms. Structures were generated using CYLview³⁷. *E*, energy.

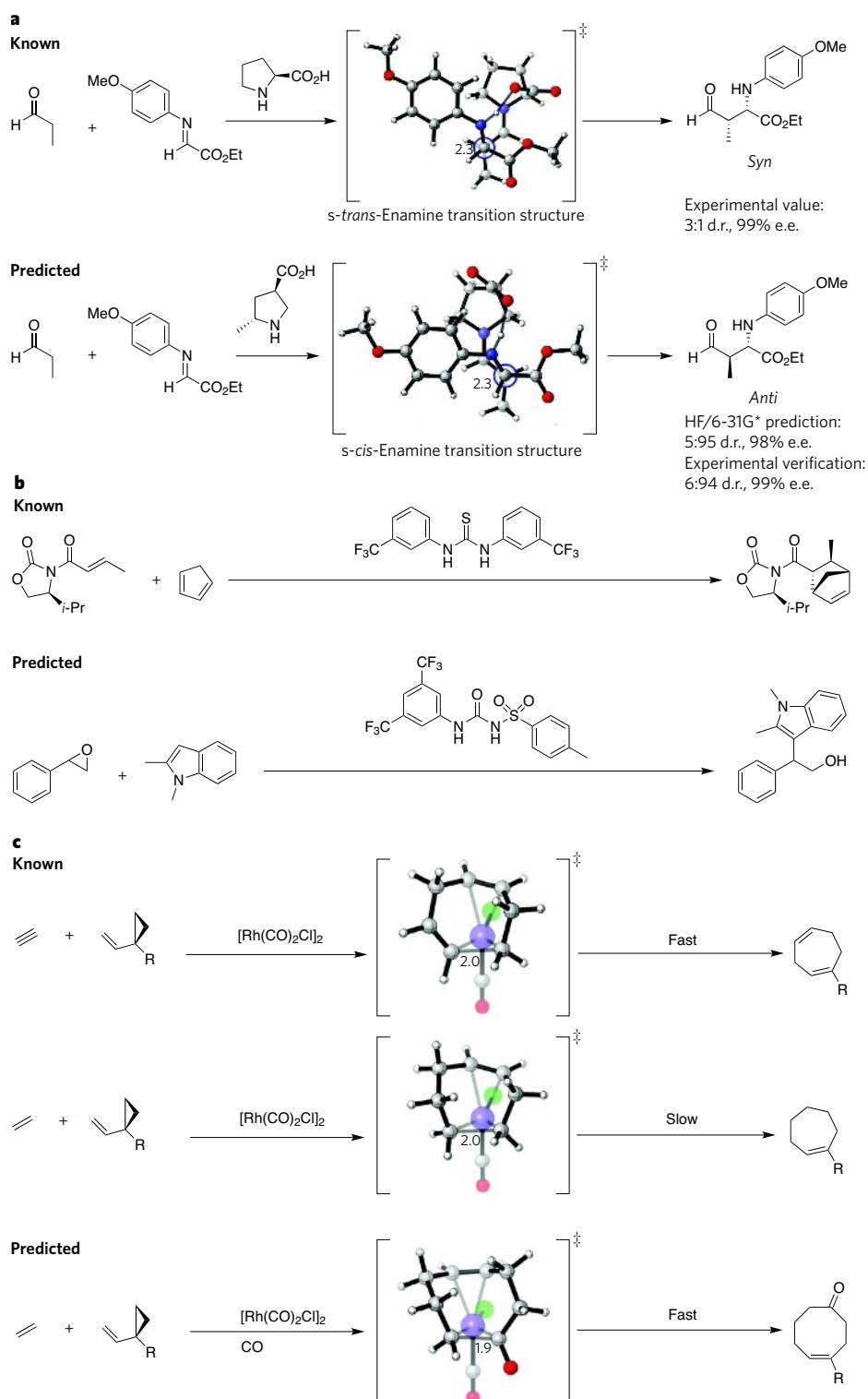


Figure 4 | Successful computational predictions of catalysts. **a**, Most Mannich reactions catalysed by organocatalysts mainly yield products in the *syn* configuration (top). Chemists designed an organocatalyst selective for the *anti* configuration (bottom) based on computational analysis of information from a proline-catalysed Mannich reaction¹⁶. **b**, The design of an epoxide-activating urea catalyst (bottom) came from the computational analysis of known thiourea catalysts (top) that activate carbonyl-containing compounds¹⁹. **c**, Rhodium-catalysed [5 + 2]-cycloaddition reactions of vinylcyclopropanes with allenes (not shown) or alkynes (top) are fast, but the analogous reactions with alkenes (centre) are slow. Computational analysis of these reactions led to the design of an analogous [5 + 2 + 1]-multicomponent-cycloaddition reaction that yields cyclooctenones²³ (bottom). Carbon is shown in grey, hydrogen in white, oxygen in red, nitrogen in blue, chloride in green and rhodium in purple (large spheres). Critical carbon–carbon bond-forming distances in the transition structure are indicated in angstroms. Transition-state structures were generated using CYLview³⁷. d.r., diastereomeric ratio; e.e., enantiomeric excess; Et, ethyl; *i*-Pr, isopropyl; R, alkoxy, alkyl, silyloxy; ‡, transition state.

computational methods. Most computational studies use polarizable continuum solvation models to compute solvation energies. These models use a polarizable continuum to compute the extent to which the solvent stabilizes the solute's dipole, as well as surface and cavitation terms to describe the solute's contact with the solvent²⁷. Specific individual interactions, such as hydrogen bonding, are overlooked in these models but can be captured by more elaborate models, such as explicit solvation dynamics²⁸. It should be noted that there are also pitfalls in using quantum-mechanical methods, such as the commonly used density-functional-theory calculations. These methods cannot accurately model dispersion forces (such as van der Waals interactions), which result from an electron-correlation effect.

Although they have not been perfected, several recent theoretical advances are leading to more efficient and more accurate computations³.

Catalysts for all occasions

What is the prospect of predicting, from scratch, the efficacy of a catalyst for any arbitrary reaction? Is it even possible to predict accurately what occurs when a mixture of chemicals is placed in a solvent? These predictions are impractical at present, because of the numerous potential reactions that typically need to be studied to define any single system. It is clear, however, that computational chemistry is suited to predicting the activation energies and stereoselectivities of particular steps in

a reaction pathway or even of whole catalytic cycles. For example, the complete catalytic cycles of proline-catalysed aldol reactions^{10,11} and of transition-metal-catalysed cycloaddition reactions^{23,29} have been studied computationally. But, in these cases, the overall reactions were known, so all of the potential competing reactions did not need to be computed.

Despite these issues, there have been some moderate successes in the *a priori* computational design of effective catalysts, such as the 'theozyme' approach. A theozyme is a theoretical catalyst, designed by computing the optimal geometry of functional groups to stabilize an uncatalysed transition state³⁰. Theozymes were originally devised to help to identify the catalytic residues of enzymes, but the same idea can be used to design catalysts. Hendrik Zipse and colleagues, for example, have used the computed transition structure for an uncatalysed aminolysis reaction³¹ to predict a host molecule that encapsulates and moderately stabilizes the transition state, leading to moderate catalysis³². Theozymes are also the starting point for the computational design of proteins that have folding properties and catalytic functions that differ from natural enzymes; two successful examples of this type have been reported, showing the power of computational design^{33,34}.

The *de novo* prediction of effective catalysts for any given reaction is a distant goal, but the computational prediction and design of more effective analogues of known catalysts is a reality. Similarly, during drug discovery, no one expects a drug to be designed from scratch using computational methods, but such techniques are invaluable for optimizing existing leads. At present, however, the use of computational methods for the design of catalysts is relatively rare, whereas it is an accepted — and some would say essential — part of drug discovery.

It is time for chemists' perceptions of computational techniques to change. These techniques provide information about known catalytic reactions that is not available from experiments alone, but this information now extends beyond that obtained by mere rationalizations. As the examples discussed here show, computational techniques have become an invaluable tool for predicting the behaviour of catalysts and have earned their place as a standard tool for the design of catalysts.

K. N. Houk and Paul Ha-Yeon Cheong are at the University of California, Los Angeles, Department of Chemistry and Biochemistry, 607 Charles E. Young Drive East, Los Angeles, California 90095, USA.

- Gaussian 03, Revision C.02 (Gaussian, Wallingford, Connecticut, 2004).
- Zhao, Y. & Truhlar, D. G. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.* **2**, 157–167 (2008).
- Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models* 2nd edn (Wiley, 2004).
- Weik, M. H. in *A Third Survey of Domestic Electronic Digital Computing Systems*. Report No. 1115, 526–536 (Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland, 1961).
- Rudolf, K., Spellmeyer, D. C. & Houk, K. N. Prediction and experimental verification of the stereoselective electrocyclization of 3-formylcyclobutene. *J. Org. Chem.* **52**, 3708–3710 (1987).
- Murakami, M., Miyamoto, Y. & Ito, Y. A silyl substituent can dictate a concerted electrocyclic pathway: inward torquoselectivity in the ring opening of 3-silyl-1-cyclobutene. *Angew. Chem. Int. Edn Engl.* **40**, 189–190 (2001).
- Murakami, M., Hasegawa, M. & Igawa, H. Theoretical and experimental studies on the thermal ring-opening reaction of cyclobutene having a stannyl substituent at the 3-position. *J. Org. Chem.* **69**, 587–590 (2004).
- Mukherjee, D., Wu, Y.-D., Fronczek, F. R. & Houk, K. N. Experimental tests of models to predict nucleophilic addition stereochemistries. *J. Am. Chem. Soc.* **110**, 3328–3330 (1998).
- Bahmanyar, S. & Houk, K. N. Transition states of amine-catalyzed aldol reactions involving enamine intermediates: theoretical studies of mechanism, reactivity, and stereoselectivity. *J. Am. Chem. Soc.* **123**, 11273–11283 (2001).
- Rankin, K. R., Gauld, J. W. & Boyd, R. J. Density functional study of the proline-catalyzed direct aldol reaction. *J. Phys. Chem. A* **106**, 5155–5159 (2002).
- Clemente, F. R. & Houk, K. N. Computational evidence for the enamine mechanism of intramolecular aldol reaction catalyzed by proline. *Angew. Chem. Int. Edn Engl.* **43**, 5766–5768 (2004).
- Bahmanyar, S. & Houk, K. N. The origin of stereoselectivity in proline-catalyzed intramolecular aldol reactions. *J. Am. Chem. Soc.* **123**, 12911–12912 (2001).
- Tang, Z. *et al.* Novel small organic molecules for a highly enantioselective direct aldol reaction. *J. Am. Chem. Soc.* **125**, 5262–5263 (2003).
- Shinisha, C. B. & Sunoj, R. B. Bicyclic proline analogues as organocatalysts for stereoselective aldol reactions: an *in silico* DFT study. *Org. Biomol. Chem.* **5**, 1287–1294 (2007).
- Cheong, P. H.-Y. & Houk, K. N. Origins of predictions of stereoselectivity in intramolecular aldol reactions catalyzed by proline derivatives. *Synthesis* **9**, 1533–1537 (2005).
- Mitsumori, S. *et al.* Direct asymmetric *anti*-Mannich-type reactions catalyzed by a designed amino acid. *J. Am. Chem. Soc.* **128**, 1040–1041 (2006).
- Cheong, P. H.-Y. *et al.* Pipecolic acid-catalyzed direct asymmetric Mannich reactions. *Org. Lett.* **8**, 811–814 (2006).
- Schreiner, P. R. & Wittkopp, A. H-bonding additives act like Lewis acid catalysts. *Org. Lett.* **4**, 217–220 (2002).
- Fleming, E. M., Quigley, C., Rozas, I. & Connon, S. J. Computational study-led organocatalyst design: a novel, highly active urea-based catalyst for addition reactions to epoxides. *J. Org. Chem.* **73**, 948–956 (2008).
- Kozlowski, M. C., Dixon, S. L., Panda, M. & Lauri, G. Quantum mechanical models correlating structure with selectivity: predicting the enantioselectivity of β -amino alcohol catalysts in aldehyde alkylation. *J. Am. Chem. Soc.* **125**, 6614–6615 (2003).
- Ianni, J. C., Annamalai, V., Phuan, P.-W. & Kozlowski, M. C. A priori theoretical prediction of selectivity in asymmetric catalysis: design of new chiral catalysts using quantum molecular interaction fields. *Angew. Chem. Int. Edn Engl.* **45**, 5502–5505 (2006).
- Yu, Z.-X. *et al.* Origins of differences in reactivities of alkenes, alkynes, and allenes in $[\text{Rh}(\text{CO})_2\text{Cl}]_2$ -catalyzed (5 + 2) cycloaddition reactions with vinylcyclopropanes. *J. Am. Chem. Soc.* **130**, 2378–2379 (2008).
- Wang, Y. *et al.* A computationally designed Rh(I)-catalyzed two-component [5 + 2 + 1] cycloaddition of ene-vinylcyclopropanes and CO for the synthesis of cyclooctenones. *J. Am. Chem. Soc.* **129**, 10060–10061 (2007).
- Corbeil, C. R., Thielges, S., Schwartzentruber, J. A. & Moitessier, N. Toward a computational tool predicting the stereochemical outcome of asymmetric reactions: development and application of a rapid and accurate program based on organic principles. *Angew. Chem. Int. Edn Engl.* **47**, 2635–2638 (2008).
- Harriman, D. J., Lambropoulos, A. & Deslongchamps, G. *In silico* correlation of enantioselectivity for the TADDOL catalyzed asymmetric hetero-Diels-Alder reaction. *Tetrahedr. Lett.* **48**, 689–692 (2007).
- Anderson, C. D., Dudding, T., Gordillo, R. & Houk, K. N. Origin of enantioselection in hetero-Diels-Alder reactions catalyzed by naphthyl-TADDOL. *Org. Lett.* **10**, 2749–2752 (2008).
- Takano, Y. & Houk, K. N. Benchmarking the conductor-like polarizable continuum model (CPCM) for aqueous solvation free energies of neutral and ionic organic molecules. *J. Chem. Theory Comput.* **1**, 70–77 (2005).
- van Erp, T. S. & Meijer, E. J. *Ab initio* molecular dynamics study of aqueous solvation of ethanol and ethylene. *J. Chem. Phys.* **118**, 8831–8840 (2003).
- Pitcock, W. H. Jr, Lord, R. L. & Baik, M.-H. The mechanism of the rhodium(I)-catalyzed [2 + 2 + 1] carbocyclization reaction of dienes and CO: a computational study. *J. Am. Chem. Soc.* **130**, 5821–5830 (2008).
- Tantillo, D. J., Chen, J. & Houk, K. N. Theozymes and compuzymes: theoretical models for biological catalysis. *Curr. Opin. Chem. Biol.* **2**, 743–750 (1998).
- Zipse, H., Wang, L.-H. & Houk, K. N. Polyether catalysis of ester aminolysis — a computational and experimental study. *Liebigs Ann.* **1996**, 1511–1522 (1996).
- Wang, L.-H. & Zipse, H. Bifunctional catalysis of ester aminolysis — a computational and experimental study. *Liebigs Ann.* **1996**, 1501–1509 (1996).
- Rothlisberger, D. *et al.* Kemp elimination catalysts by computational enzyme design. *Nature* **453**, 190–195 (2008).
- Jiang, L. *et al.* *De novo* computational design of retro-aldol enzymes. *Science* **319**, 1387–1391 (2008).
- The PyMOL Molecular Graphics System v1.1 (DeLano Scientific, Palo Alto, 2007).
- Tarini, M., Cignoni, P. & Montani, C. Ambient occlusion and edge cueing to enhance real time molecular visualization. *IEEE Trans. Vis. Comput. Graph.* **12**, 1237–1244 (2006).
- CYLview v1.0b (C. Legault, 2007).

Acknowledgements We are grateful to the University of California, Los Angeles, as well as to the National Institute of General Medical Sciences (grant GM 36700) and the National Science Foundation (grant CHE-0548209), for financial support. We also thank C. Legault for the use of CYLview, and M. Tarini, P. Cignoni and C. Montani for assistance in using QuteMol.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to K.N.H. (hok@chem.ucla.edu).

Carbon-heteroatom bond formation catalysed by organometallic complexes

John F. Hartwig¹

At one time the synthetic chemist's last resort, reactions catalysed by transition metals are now the preferred method for synthesizing many types of organic molecule. A recent success in this type of catalysis is the discovery of reactions that form bonds between carbon and heteroatoms (such as nitrogen, oxygen, sulphur, silicon and boron) via complexes of transition metals with amides, alkoxides, thiolates, silyl groups or boryl groups. The development of these catalytic processes has been supported by the discovery of new elementary reactions that occur at metal-heteroatom bonds and by the identification of factors that control these reactions. Together, these findings have led to new synthetic processes that are in daily use and have formed a foundation for the development of processes that are likely to be central to synthetic chemistry in the future.

Organometallic complexes contain organic groups attached to a central metal through metal-carbon (M-C) bonds. These complexes undergo a set of elementary reactions, many of which release products containing new C-C or C-H bonds. The linking of these reactions to generate catalytic processes that form C-C bonds has been a recent focus of synthetic chemistry research. Such organometallic catalysts are now used to form C-C bonds in commodity chemicals (cheap chemicals sold in bulk) and polymers that are produced on the scale of millions of tonnes per year. They are also used in tailor-made polymers that are produced in small quantities and in highly intricate pharmaceuticals and biologically active natural products (see page 323), which are typically produced in milligram quantities for biological testing.

The backbone of many organic compounds is composed of C-C bonds, but the function of these molecules is often derived from the presence of heteroatoms, such as nitrogen, oxygen and sulphur, which are held in these molecules by C-heteroatom bonds. For example, pharmaceuticals and conductive polymers (plastics that conduct electricity) often contain amine C-N bonds, and almost all natural products contain ether, ketone or ester C-O bonds. Heterocyclic compounds in which C-N, C-O or C-S bonds are present in the ring structure are found in all applications of chemistry. Moreover, useful intermediates in synthesis often contain C-B or C-Si bonds that are later converted into C-C, C-O or C-N bonds in the final products.

Amines, alkoxides, thiolates and related reagents are common nucleophiles in reactions that occur without a catalyst. However, these reagents do not react with the weak electrophiles that are useful for synthesis, such as aromatic halides, and they do not react with common electron-rich reagents, such as alkenes. Thus, the ability to carry out catalytic reactions that form these C-heteroatom bonds would substantially affect the synthesis of molecules with important functions. Such catalytic reactions would involve reactions at M-heteroatom bonds that resemble classic organometallic reactions at M-C bonds. Compounds with M-heteroatom bonds that undergo these types of reaction are referred to here as hetero-organometallic complexes. To tap the potential of these compounds as intermediates in emerging catalytic processes, the principles that control their reactivity need to be understood, and these are now beginning to be delineated.

In this review, I describe carbon-heteroatom bond formation catalysed by organometallic complexes and the underlying hetero-organometallic chemistry. I focus on several examples: cross-coupling to form C-N, C-O and C-S bonds¹⁻³; C-H bond functionalization to form C-O (ref. 4), C-X (where X is a halide) (refs 5, 6) and C-B (ref. 7) bonds; and olefin oxidations⁸ and aminations^{9,10} to form C-O and C-N bonds (where olefins are compounds that contain at least one C-C double bond, and alkenes are aliphatic olefins with the general formula C_nH_{2n}). In addition, the principles of metal-ligand bonding are used as a framework to explain the differences in reactivity between organometallic intermediates and hetero-organometallic intermediates.

Classic processes catalysed by organometallic systems

Several classic catalytic reactions of organometallic systems have been used for decades. Three of these processes — cross-coupling, hydrogenations of olefins and metathesis of olefins (Fig. 1a-c) — are commonly used by synthetic chemists. A fourth process — dehydrogenation of alkanes (Fig. 1d) — has been highly sought and is being developed.

Palladium-catalysed cross-coupling to form C-C bonds is one of the most frequently used catalytic processes for synthesizing medicinally active compounds. Examples of products prepared by this method are the antihypertensive drugs known as sartans¹¹, the anti-asthma drug Singulair (montelukast sodium)¹², products such as sunscreen¹² and the polymers used for light-emitting diodes and for sensing the explosive trinitrotoluene (TNT)¹³⁻¹⁶. The intermediates that form the new C-C bond in the final product are classic organometallic species.

Hydrogenations of olefins are carried out on a large scale for the synthesis of commodity chemicals and on a small scale for the synthesis of fine chemicals and advanced pharmaceutical intermediates. Materials ranging from margarine (partially hydrogenated vegetable oil)¹⁷ to the most structurally and stereochemically intricate natural products are prepared by hydrogenation, and in 2001 the Nobel Prize in Chemistry was awarded in part for work on enantioselective hydrogenation¹⁸.

Another classic catalytic reaction of olefins — the cleavage and re-formation of C=C bonds, known as olefin metathesis — occurs through complexes that contain M=C bonds, and the chemists behind

¹Department of Chemistry, University of Illinois, 600 South Mathews Avenue, Urbana, Illinois 61801, USA.

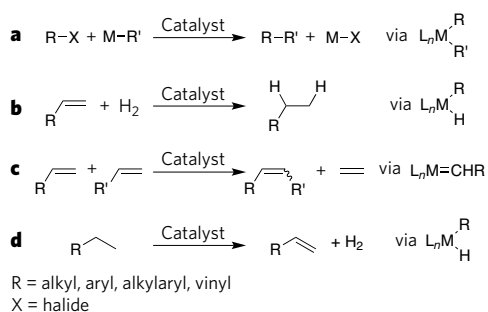


Figure 1 | Catalytic C–C and C–H bond-forming processes. Cross-coupling (a), hydrogenation of olefins (b) and metathesis of olefins (c) are commonly used reactions in organic synthesis. Alkane dehydrogenation (d) has been studied as a means to modify typically unreactive alkanes and has been combined with alkene metathesis to develop a new alkane metathesis. L_n , ancillary ligand; M, metal.

this work were awarded the Nobel Prize in Chemistry in 2005 (see ref. 19 for a recent review).

The conversion of typically unreactive C–H bonds to C=C and C–X bonds (where X is O, N, B or Si)^{20,21} holds promise for reducing the reliance on existing functional groups. The cleavage of two C–H bonds in an alkane to form the C=C π -bonds in alkenes (Fig. 1d) has been studied intensively. Such dehydrogenation has recently been used, in tandem with olefin metathesis, to develop a catalytic process termed alkane metathesis²², which cleaves and reforms the C–C bonds in alkanes.

Classic elementary reactions of organometallic complexes

Catalytic processes consist of a series of elementary, stoichiometric reactions. Advances in the synthesis of transition-metal complexes, the discovery of new elementary, stoichiometric reactions of these complexes, and the development of methods to link these steps together into catalytic cycles have led to the development of many of the catalytic organometallic processes used today. Typical elementary reactions of organometallic complexes in such catalytic cycles include oxidative addition, reductive elimination, migratory insertion, β -hydrogen elimination and [2 + 2] cycloadditions^{23,24} (Fig. 2). Oxidative addition adds an organic reagent to a transition metal through insertion of the metal into one of the bonds of the incoming reagent. Reductive elimination extrudes an organic product by coupling two ligands on the metal. Migratory insertion leads to the incorporation of a bound, neutral (dative) ligand into a metal–ligand covalent bond, and β -elimination extrudes such a ligand.

These elementary reactions (and several others) have been used to create hundreds of catalytic processes. Cross-coupling (Fig. 2a) occurs by a sequence of steps initiated by the oxidative addition of an organic halide and finished by reductive elimination. Hydrogenation (Fig. 2b) often occurs by a combination of oxidative addition of the H–H bond in dihydrogen, olefin insertion to form the first C–H bond and reductive elimination to form the second C–H bond. The closely related and industrially important addition of silanes to alkenes occurs by a similar sequence that is initiated by oxidative addition of an H–Si bond²⁵, and many reactions that lead to functionalization of C–H bonds begin with oxidative addition of a C–H bond. The metathesis of olefins occurs by a series of [2 + 2] and retro [2 + 2] cycloadditions (Fig. 2c).

Elementary reactions of hetero-organometallic complexes

Although the classic organometallic reactions, which form C–H bonds and C–C bonds, have been known for decades, few examples of reactions that form or cleave M–heteroatom bonds (shown generically in Fig. 3) have been recorded until fairly recently. The absence of this reaction chemistry hampered efforts to develop catalytic cross-couplings, additions to olefins, and C–H bond functionalizations that form C–N, C–O or C–S bonds. Until recently, there were no isolated transition-metal complexes that underwent reductive elimination to form C–N, C–O and C–S bonds in amines, ethers and sulphides^{26,27}, and few compounds reacted

with the N–H bond in ammonia to form a monomeric product^{28,29}. Isolated complexes that inserted simple alkenes into the M–N or M–O bonds of M–amido or M–alkoxo complexes were also unknown^{30–32}. And M–imido or M–oxo complexes that undergo [2 + 2] cycloadditions with alkenes (in which the two atoms of the M=O or M=N π -bond combine to make a saturated four-membered ring) were not known until the past decade³³. The limited precedent for these processes raised the question of whether it was possible to develop organometallic processes to form amines, ethers and sulphides, or whether the metal–ligand combinations necessary to trigger this reactivity had not yet been identified. Recent progress clearly indicates that the latter was true.

Much of the literature on M–alkoxo or M–amido complexes can be divided into two sets. In one set, high-valent early-transition M–alkoxo and M–amido complexes were commonly used, but the amides and alkoxides were typically ancillary ligands because the M–O and M–N bonds in these complexes were too strong to display extensive reactivity. In the second set, methods to prepare late-transition-metal–amido and –alkoxo complexes were beginning to be developed, and the properties of these complexes were beginning to be explored³⁴. These complexes, however, were often too unstable towards β -hydrogen elimination to observe reactions that would form C–N or C–O bonds.

In recent years, our understanding of the reactions of M–amido, M–alkoxo and M–thiolato complexes has changed markedly. Many researchers are beginning to develop synthetically valuable processes relying on hetero-organometallic intermediates. One current fundamental issue is how the differences in properties between carbon and a heteroatom change the course of these reactions. For example, does an

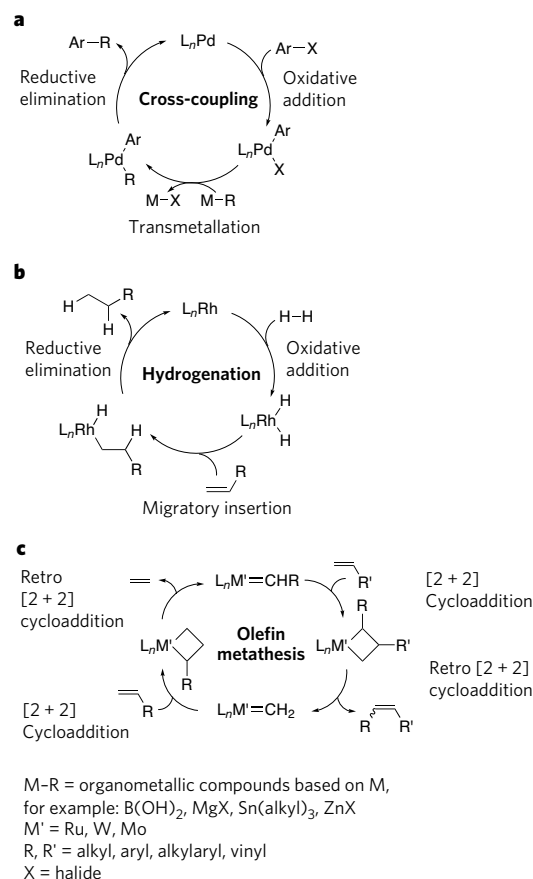


Figure 2 | Mechanisms of three common catalytic organometallic processes. **a**, Cross-coupling is initiated by oxidative addition of an organic halide, continues through transmetalation and terminates by reductive elimination. **b**, Hydrogenation begins with oxidative addition of the H–H bond in dihydrogen, migratory insertion of an olefin to form a C–H bond and reductive elimination to form a second C–H bond. **c**, The metathesis of olefins occurs through a series of [2 + 2] cycloadditions and retro [2 + 2] cycloadditions. Ar, aryl.

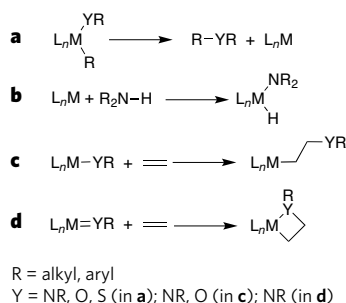


Figure 3 | Recently discovered organometallic reactions of transition-M-heteroatom bonds. **a**, Reductive elimination to form C–N, C–O and C–S bonds in amines, ethers and thioethers. **b**, Oxidative addition of amine N–H bonds. **c**, Migratory insertions of olefins into metal amides and metal alkoxides. **d**, [2 + 2] Cycloadditions between olefins and M–imido or M–oxo complexes. These reactions are analogues of classic reactions occurring at M–C bonds and have only recently been discovered.

increase in the electronegativity of the atom bound to the metal increase or decrease the rate of addition, insertion, elimination or cycloaddition? How does the presence of an electron pair on the heteroatom affect the rates of these reactions? Does the presence of an unoccupied valence orbital on the heteroatom affect the rates of these reactions? The examples described here show that the set of guidelines needed to explain the patterns of reactivity is more complex than matching or mismatching hard or soft ligands with hard or soft metals³⁵. A few such guidelines are presented in the context of the emerging catalytic processes that form C–heteroatom bonds through the reactions of hetero-organometallic species: namely cross-coupling to form C–N, C–O and C–S bonds; C–H bond functionalization to form C–O, C–X (where X is a halide) and C–B bonds; and olefin oxidations and aminations to form C–O and C–N bonds.

Catalysis by carbon–heteroatom reductive elimination

Cross-coupling reactions to form the C–N, C–O and C–S bonds in amines, ethers and sulphides (Fig. 4a) are widely used for the synthesis of pharmaceutical candidates, fine chemicals, polymers, components of organic devices and even ligands for other catalysts^{1,36}. The C–N coupling process evolved from an initially³⁷ promising, but relatively impractical, coupling of tin amides with aryl halides in the presence of a palladium catalyst containing a sterically hindered, monodentate (bonded through a single atom) aromatic phosphine³⁸ into a general process that has been made practical by several successive generations of catalysts.

The initial process involving tin amides was replaced by reactions involving amines and a base^{39,40}. The scope of the new process became broader with the use of bidentate aromatic phosphines that inhibited competing β -hydrogen elimination^{41,42}, and later with the use of sterically hindered alkyl monophosphines^{1,43–48} that allowed the activation of less reactive haloarenes and the coupling of weaker oxygen nucleophiles. Recent ‘fourth-generation’ catalysts containing sterically hindered alkyl bisphosphines further improved catalyst efficiency^{49,50}. Overall, the development of these processes began to demonstrate the capability of late-transition-metal–amido, –alkoxo and –thiolato complexes to participate productively in catalytic cycles.

The basic steps of the mechanism of the amination process are shown in Fig. 4b. Analogous steps lead to catalytic cycles to form ethers and thioethers. Like cross-coupling to form C–C bonds, this process is initiated by oxidative addition of a haloarene. An arylpalladium–amido, –alkoxo or –thiolato complex is then formed from the oxidative addition product by reaction of an amine, alcohol, or thiol and base. The catalytic cycle is then completed by reductive elimination to form the C–N, C–O or C–S bond in the product amine, ether or thioether.

The scope of the final reductive elimination reaction in this catalytic cycle was striking. Although reductive eliminations to form the C–N, C–O or C–S bond in an amine, ether or thioether were unknown when

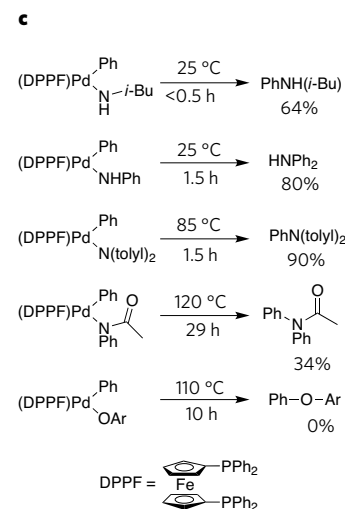
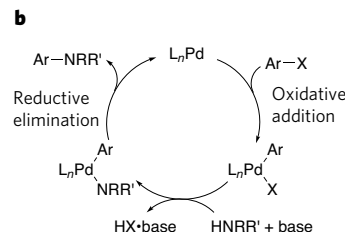
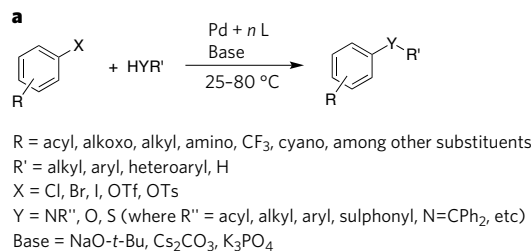


Figure 4 | Aspects of palladium-catalysed amination of aryl halides and related reactions. **a**, The overall transformation in the presence of a palladium (Pd) catalyst. Four generations of ligands (L) have been developed for the catalytic process. The first-generation catalyst contained P(*o*-tolyl)₃. The second-generation catalysts contained chelating aromatic phosphines as ligand: DPPF (1,1'-bis(diphenylphosphino)ferrocene), BINAP (2,2'-bis(diphenylphosphino)-1,1'-binaphthyl) and xantphos (4,5-bis(diphenylphosphino)-9,9-dimethylxanthene). The third-generation catalysts contain hindered alkylphosphines and carbenes: P(*t*-Bu)₃, Ph₃FCP(*t*-Bu)₂, *N*-heterocyclic carbenes, (biaryl)PR₂, (heterobiaryl)PR₂ and caged P(NRR')₃. The fourth (most recent) generation of catalysts contain hindered ferrocenyl alkyl bisphosphines. **b**, The catalytic cycle involving oxidative addition of an aryl halide, formation of an arylpalladium amide and reductive elimination of an amine. **c**, Reactions revealing the electronic effects on reductive elimination to form C–N bonds in amines and amides, and C–O bonds in ethers. The rates of reductive elimination from Pd(II) are faster when the heteroatom bound to the metal is more electron rich. Fc, ferrocenyl; *i*-Bu, isobutyl; Me, methyl; OTf, trifluoromethanesulphonate; OTs, 4-toluenesulphonate; Ph, phenyl; *t*-Bu, *tert*-butyl; tolyl, C₆H₄-4-Me.

the coupling reactions were first being developed, the broad scope of catalytic couplings of aryl halides with amines, alcohols and thiols now implies that these types of reductive elimination can occur from arylpalladium(II) complexes containing a diverse set of aryl and heteroaryl groups, as well as a diverse set of amido, alkoxo and thiolato groups. In fact, with the right ancillary ligands on the metal, these reactions are typically faster than competing processes, such as β -hydrogen eliminations, that lead to undesired side products.

Electronic effects on carbon–heteroatom reductive elimination

Although amido and alkoxo ligands have now been shown to participate in the type of reductive elimination reactions that alkyl ligands undergo, the electronic properties of amido and alkoxo ligands cause the rates and scope of these reactions to differ in synthetically important ways²⁷.

The reactions^{51,52} in Fig. 4c show that the rate of reductive elimination from a series of compounds containing the same ancillary ligand is faster when the covalent heteroatom ligand has stronger electron-donating properties. Comparisons of the rates of reactions of arylamido complexes and phenoxo complexes⁵³, and the rates of reaction of alkylamido versus alkoxo complexes⁵⁴, also show that the rate of reductive elimination from the amido complexes is faster than from complexes containing less electron-donating phenoxo and alkoxo ligands. But a comparison of the rates of these reactions to those of thiolate complexes shows that basicity alone does not control the rate. Complexes containing the less basic, but more polarizable and nucleophilic, thiolato ligand undergo reductive elimination much faster than do alkoxo complexes and at rates that are similar to, or faster than, those of the amido complexes⁵⁵.

These relative rates clearly argue against a trend in relative reactivity based on hard–soft matches and mismatches. The compounds containing the largest hard–soft mismatch are the least reactive. (An amidate ligand is harder than an amide ligand, and a phenoxide is harder than an arylamide.) Two possible alternative explanations could be based on the participation of the electron pair on the heteroatom. In one case, the compounds containing the most basic electron pairs would be the most reactive because these electron pairs weaken the metal–ligand bond through filled–filled $d\pi$ – $p\pi$ -orbital interactions⁵⁶; in a second case, the reaction would occur by attack of the electron pair on the palladium-bound aryl group, and the most nucleophilic electron pair would lead to the fastest rate. If either of these theories is correct, then substantially smaller electronic effects would be observed for analogous reactions of alkyl complexes because the alkyl complexes lack the basic electron pair on the atom bound to the metal.

Studies on reductive elimination from related arylpalladium–alkyl complexes in which the alkyl groups have varied electronic properties⁵⁷ showed that the electronic effects on reductive elimination from the alkyl complexes were similar to those on reductive elimination from the amido complexes. Thus, the relative rates for reductive elimination from amido complexes are unlikely to result from the properties of the lone pair on nitrogen. Instead, the dominant electronic effect on reductive elimination seems to result from differences in the electronegativity of the atom bound to the metal⁵⁸. Apparently, the more covalent and less ionic the M–heteroatom bond, and the more polarizable the heteroatom, the faster the rate of concerted reductive elimination from the intermediates in the catalytic coupling to form C–heteroatom bonds^{59,60}.

Recent studies on catalytic coupling and stoichiometric reductive elimination from platinum(IV) have shown that the metal and the reaction mechanism can change the electronic effects on reductive elimination. Reductive eliminations of alkyl aryl ethers, alkyl acetates and *N*-alkyl sulphonamides from platinum(IV) that are part of the mechanism of alkane functionalizations (discussed in the next section) occur by a non-concerted mechanism, and these reactions occur faster from complexes containing ligands that can better support anionic charge at the heteroatom than from those whose ligands are less able to support such charge^{59,60}. Reductive elimination from arylpalladium(IV) complexes has also been observed, but it is too soon to draw conclusions about the electronic effects on these reductive eliminations⁶¹. In addition, catalytic couplings that form C–N and C–O bonds in the presence of iron^{62,63} and copper⁶⁴ catalysts are being developed, and future studies on the intermediates in these processes should show whether the electronic effects on reductive elimination from these transition metals are more like those on reductive elimination from palladium(II) or platinum(IV).

Catalytic C–H bond functionalization

In addition to cross-couplings that occur by reductive elimination to form C–heteroatom bonds, catalytic C–H bond functionalizations that occur through this elementary reaction have been developed. Two

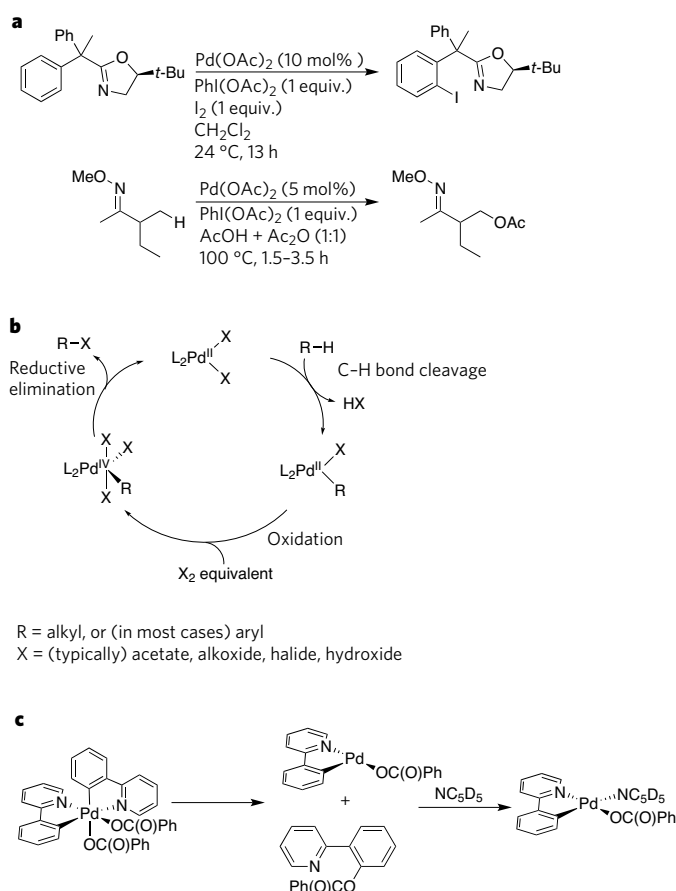


Figure 5 | Organometallic oxidative C–O and C–halogen bond-forming functionalization of C–H bonds. **a**, Two representative directed functionalizations. In the upper example, the metal catalyst binds to the nitrogen of the oxazoline substituent, and this binding positions the catalyst for cleavage and functionalization of the *ortho* C–H bond on the phenyl ring. The combination of the $PhI(OAc)_2$ oxidant and iodine is thought to lead to an intermediate arylpalladium(IV)–iodide complex, which forms the aryl C–I bond in the product. In the lower example, a related process occurs. In this case, binding of the catalyst to the oxime nitrogen positions it for cleavage of the C–H bond shown. The $PhI(OAc)_2$ oxidant is then thought to lead to an alkylpalladium(IV) acetate intermediate, which forms the C–O bond in the product. **b**, General catalytic cycle for palladium-catalysed oxidation of C–H bonds involving cleavage of a C–H bond, oxidation to a high-valent metal centre and reductive elimination to form a C–heteroatom bond. **c**, Example of the C–X bond-forming step in a well-characterized system. This reaction is thought to model the C–heteroatom bond-forming step of directed C–H-bond functionalizations. Reductive elimination to form the C–O bond on the organic product is followed by coordination of deuterated pyridine solvent. D, deuterium. equiv., equivalents; OAc, acetate.

classes of such reactions are shown in Figs 5 and 6. In one case (Fig. 5), the hydrogen in a C–H bond of an arene or alkyl chain is converted to a halide or acetate with regioselectivity controlled by a ligating functionality. This chemistry draws parallels with previous functionalization of alkyl C–H bonds catalysed by platinum complexes, originally discovered by Alexander Shilov and studied intensely in recent years with the goal of developing selective oxidation of methane^{65–68}. In a second case (Fig. 6), a terminal C–H bond in an alkane is functionalized by the use of a rhodium catalyst and boron reagents.

These processes require transition-metal complexes that can both cleave a C–H bond and form a C–heteroatom bond. Some catalysts, such as iron–oxo complexes, do so without formation of an intermediate containing an M–C bond⁶⁹. These complexes often abstract C–H bonds and deliver the resultant hydroxo group to an alkyl radical. These ‘outer-sphere’ reactions typically favour cleavage of the weaker

of the available C–H bonds. In other cases, a transition-metal complex cleaves a C–H bond to form an organometallic intermediate containing a metal–carbon bond. These ‘inner-sphere’ processes preferentially lead to cleavage of primary alkyl C–H bonds over weaker secondary alkyl C–H bonds to form *n*-alkyl intermediates^{70–72}, and preferentially add even stronger aromatic C–H bonds over alkyl C–H bonds to form M–aryl intermediates⁷¹. Thus, a method to couple the resultant organometallic ligand to an alkoxide, amide or halide ligand would create methods for inner-sphere, organometallic C–H bond ‘functionalization’ that complement the regioselectivity of methods based on outer-sphere C–H activation.

The reactions shown in Figs 5 and 6 demonstrate two approaches to C–H bond functionalization by C–heteroatom bond-forming reductive elimination. In the first approach, the barrier to reductive elimination was reduced by generating high-valent intermediates. Although reductive eliminations to form C–O and C–halogen bonds from palladium(II) intermediates are slow and require particular ligands⁵⁴, reductive eliminations to form C–O and C–halogen bonds from palladium(IV) could be more rapid and more general. Thus, oxidation of an alkylpalladium(II) product of C–H activation could generate a high-valent intermediate that would result in functionalization of the alkyl intermediate by reductive elimination to form C–O or C–halogen bonds^{61,73}. This approach, in combination with the incorporation of a functional group in the substrate that can ligate the metal, has been used to develop directed functionalization of both aromatic⁷⁴ and aliphatic C–H bonds^{4,75} (Fig. 5a).

The mechanism of the palladium-catalysed oxidations (Fig. 5b) is initiated by cleavage of the C–H bond by palladium(II) complexes. This cleavage has been shown by several studies to occur by coordination of the C–H bond to the metal, followed by deprotonation⁷⁶. Oxidation of the palladium(II) intermediate then forms a palladium(IV) intermediate that undergoes reductive elimination to form the C–heteroatom bond. Reductive elimination to form a C–O bond from an isolated palladium(IV) complex has now been observed directly⁶¹ (Fig. 5c).

In the second approach, the functionalization of terminal alkyl C–H bonds and sterically accessible aryl C–H bonds⁷⁷ was developed by exploiting the electronic properties of intermediates containing M–B bonds⁷ (Fig. 6a). The catalytic cycle for the alkane functionalization (Fig. 6b) involves the reaction of rhodium–boryl intermediates with primary alkyl C–H bonds to form organometallic intermediates that undergo B–C bond-forming reductive elimination. Currently available mechanistic data imply that the C–H bond cleavage and the B–C bond formation occur with low barriers, owing to participation of the unoccupied orbital on boron.

A detailed mechanism of the C–H bond cleavage step of this process⁷⁸, as deduced by calculations using density functional theory (DFT), is shown in Fig. 6c. These calculations imply that the hydrogen atom is passed from the coordinated alkane to the unoccupied *p* orbital at boron to cleave the C–H bond. These modes of C–H bond cleavage are new and result from the presence of an unoccupied *p* orbital on boron. The current computational data on this process do not address the issue of regioselectivity, but previous studies on the stoichiometric cleavage of C–H bonds to form M–alkyl products also selectively cleave terminal C–H bonds, in part owing to steric effects⁷².

This new pathway for C–H bond cleavage would not be productive without accompanying B–C bond formation. Fortunately, this B–C bond formation occurs with a low barrier, presumably because of a favourable match between the electrophilic boryl group and a nucleophilic alkyl group. Thus, the electronic properties of boron are intimately involved in both the activation and functionalization stages of this catalytic, organometallic alkane functionalization.

Catalysis by olefin insertions into M–O and M–N bonds

Addition reactions to olefins constitute some of the most desirable, and when successfully developed, most used catalytic reactions in the chemical industry. Thus, much recent effort has been focused on reactions that form C–heteroatom bonds using olefins as reagents. Several

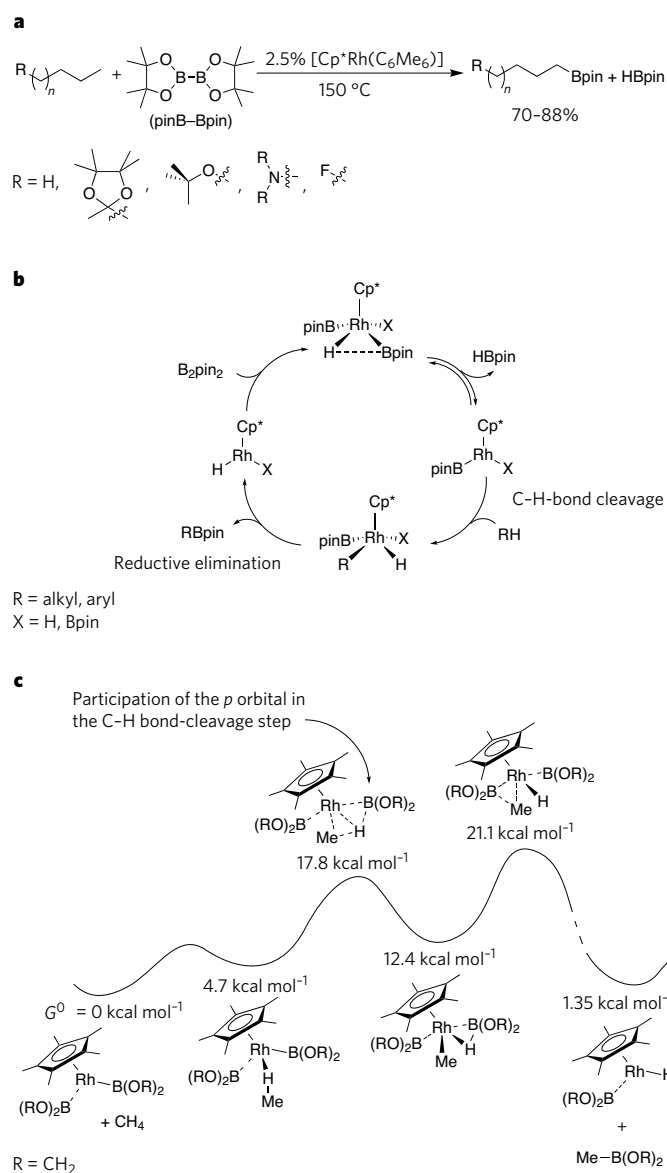


Figure 6 | Summary of the functionalization of alkanes with M–B intermediates. **a**, The overall transformation. In this process, the rhodium catalyst selectively cleaves the terminal C–H bond of an alkane or of the alkyl group of an acetal, ether, amine or alkyl fluoride, and it delivers a boryl group from the diboron (pinB–Bpin) reagent to the resultant alkyl intermediate to form the functionalized product. **b**, The catalytic cycle for alkane borylation involves C–H bond cleavage by an M–B intermediate followed by reductive elimination to form a B–C bond. **c**, The energetics of the reaction coordinate for C–H bond functionalization by M–B complexes deduced by using a density-functional-theory method. This reaction coordinate shows that the *p* orbital on boron assists in the C–H bond-cleavage step. Cp*, pentamethylcyclopentadienyl; pin, pinacolate.

types of catalyst for the oxidations and oxidative aminations of olefins through organometallic intermediates have been shown recently to react through yet another set of elementary reactions of alkoxo and amido complexes, paralleling the reactions of alkyl complexes. These reactions are olefin insertions into M–N and M–O bonds, and catalytic reactions that involve this class of elementary reaction include the following: palladium-catalysed oxidations of olefins to form aldehydes, vinyl ethers and vinyl acetates; related oxidative aminations to form enamides; and hydroaminations of alkenes to form alkylamines. Other catalysts for the hydroamination of alkenes have been shown to react through organometallic amido intermediates generated by [2 + 2] cycloadditions between imido complexes and alkenes.

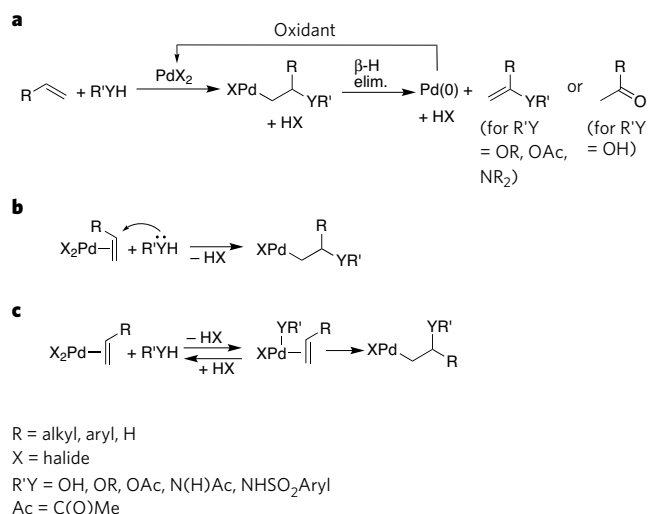


Figure 7 | Organometallic oxidation and oxidative amination of olefins.

a, The overall mechanism for the palladium-catalysed oxidation of olefins. **b**, **c**, Mechanisms for the C–heteroatom bond-forming step of the catalytic process. This step can occur by nucleophilic attack on a coordinated olefin (**b**) or migratory insertion into an M–heteroatom bond (**c**). β -H elim., β -hydrogen elimination.

Organometallic olefin oxidation and oxidative amination

A generic scheme for palladium-catalysed olefin oxidation is shown in Fig. 7. Extensive mechanistic studies on this process have led researchers to agree on the intermediacy of hydroxyalkyl, alkoxyalkyl, acetoxyalkyl and aminoalkyl complexes (Fig. 7a). However, there has been a longstanding debate about the mechanism for formation of these intermediates⁷⁹. By one mechanism (Fig. 7b), the β -functionalized alkyl intermediate is formed by coordination of an olefin and subsequent nucleophilic attack onto the coordinated olefin. By an alternative mechanism (Fig. 7c), the olefin coordinates and then inserts into an accompanying hydroxo, alkoxy, carboxylate or amidate ligand. Several recent experiments have shown that insertion of the olefin into an M–O or M–N bond leads to hydroxyalkyl, alkoxyalkyl or aminoalkyl intermediates under some of the commonly used reaction conditions.

The stereochemistry of certain oxidation products differentiates these two paths. External nucleophilic attack leads to an *anti* addition of the metal and the nucleophile across the C=C bond, whereas insertion leads to a *syn* addition of the two. Oxidations of internal olefins or deuterium-labelled terminal olefins in the presence of a source of electrophilic chlorine (such as CuCl₂) would lead to different stereoisomeric halohydrin products from the two mechanisms. The identity of products from cyclic substrates can also reveal the mechanism of formation of the hydroxyalkyl, alkoxyalkyl or aminoalkyl intermediate.

These studies^{80–84} showed that reactions conducted in the absence of chloride form products through a *syn* addition across the C=C bond, whereas analogous reactions conducted in the presence of chloride form products through an *anti* addition across the C=C bond. Most probably, at high concentrations of added chloride, coordination of chloride occurs and discourages the generation of a hydroxo or alkoxy ligand positioned *cis* to the olefin, whereas in the absence of chloride, hydroxo or alkoxy ligands are generated *cis* to the olefin, and subsequent migratory insertion occurs. Recent related experiments on the oxidative amination of olefins provide similar stereochemical evidence for migratory insertion of an olefin into a palladium amidate bond⁹.

Hydroaminations of olefins through metal-amido complexes

Some of the catalysts for olefin hydroamination have also been shown to occur by insertion of an olefin into an M–amide linkage. Complexes that contain two substituted cyclopentadienyl ligands on lanthanide metals were some of the first catalysts for intramolecular hydroaminations of alkenes to form nitrogen heterocycles that occurred with substantial

turnover numbers⁸⁵. These reactions have been recently extended to the formation of long-chain amines by the combination of ethylene and a dialkylamine⁸⁶. The initial work has led to a family of lanthanide complexes that catalyse intramolecular alkene hydroamination^{87–89}.

The discovery of these lanthanide-catalysed hydroaminations dispelled the idea that M–N bonds involving high-valent electrophilic metals are too strong to undergo processes that would parallel the reactions of high-valent organometallic complexes, such as olefin insertion. Measurements of lanthanide–amide and lanthanide–alkyl bond strengths led to the prediction that insertions of alkenes into lanthanide–amide bonds would be close to thermoneutral⁹⁰, and Tobin Marks first showed that lanthanide-catalysed hydroaminations of aminoalkenes occur by intramolecular insertion of olefins into lanthanide amides⁹¹ (Fig. 8a). The catalytic reactions occur over the course of hours at room temperature by rate-limiting insertion of the olefin into the lanthanide–amide bond.

An alternative approach to the hydroamination of C–C multiple bonds exploits the ability of zirconium(IV) and titanium(IV) complexes containing M–N multiple bonds (imido complexes) to undergo [2 + 2] cycloadditions with substrates containing C–C multiple bonds⁹². The catalytic cycle for the hydroamination of alkenes by neutral titanium and zirconium complexes (Fig. 8b) occurs by a [2 + 2] cycloaddition between an alkene, alkyne or allene and an M–N multiple bond to form a four-membered metallocycle containing one M–N and one M–C bond. This step is similar to that of the [2 + 2] cycloadditions between alkenes and complexes containing M=C bonds (carbene complexes) that occur during olefin metathesis. Protonation of the M–C bond in the resultant metallocycle by an amine, followed by elimination of the organic addition product regenerates the M–imido complex and releases the hydroamination product. In the past few years, ancillary ligands on the metal centre have been identified that extend these original hydroaminations of alkynes⁹² to the intramolecular hydroamination of aminoalkenes to form pyrrolidines and piperidines^{93,94}.

Insertions of olefins into metal–amide and metal–alkyl bonds

The conclusion that some oxidative aminations and hydroaminations of olefins occur by insertion of an olefin into an M–amide linkage begins

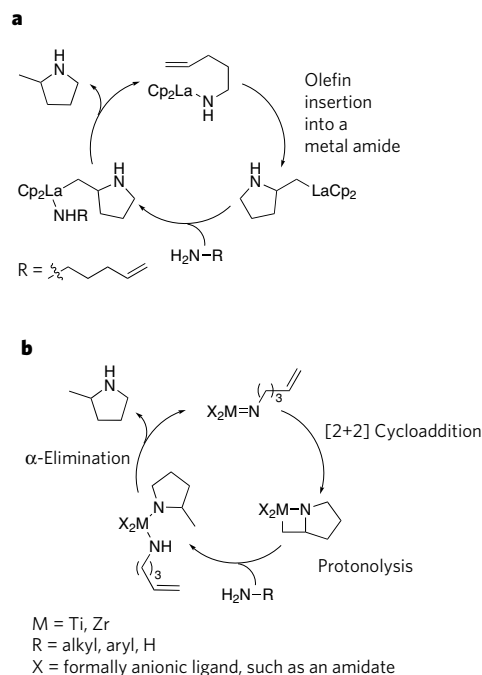


Figure 8 | Two mechanisms for catalytic hydroaminations of C–C multiple bonds through M–amido intermediates. **a**, Hydroamination of olefins catalysed by lanthanide–amides occurring by intramolecular insertions of alkenes into an M–amide bond. **b**, Hydroamination of olefins catalysed by group IV complexes occurring by [2 + 2] cycloaddition to form a metallocyclic amido intermediate. Cp, cyclopentadienyl; La, lanthanide.

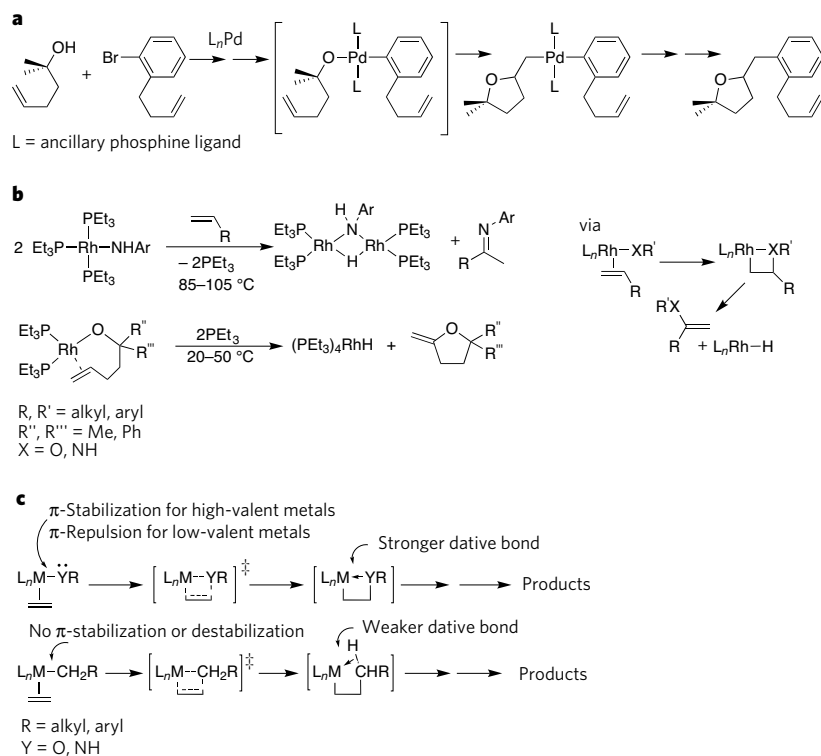


Figure 9 | Data and basis for relative rates of olefin insertion into alkyl, amide and alkoxo complexes.

a, b, Examples of insertions into late-transition-metal-alkoxo and -amido complexes. **c,** Rationalization for why the rates of olefin insertion into late-transition-metal alkoxides and amides are faster than into late-transition-metal alkyls, based on the destabilization of the alkoxo and amide reactants and the stabilization of the products of insertion into the alkoxo and amido complexes by an M–Y dative bond. Et, ethyl.

to allow a comparison of the rates of insertions of alkenes into various M–alkyl and M–amido complexes and an understanding of the factors that control these relative rates. These data provide qualitative comparisons that can be used to formulate preliminary theories about the effect of the non-bonded electron pair on rates of olefin insertions into high-valent and low-valent amido complexes.

A comparison of lanthanide-catalysed hydroamination and olefin polymerization reveals the relative rates of insertions of olefins into lanthanide-alkyl and -amide bonds. The polymerizations of alkenes by lanthanide catalysts involves intermolecular, rate-limiting insertions of ethylene into lanthanocene alkyl intermediates⁹⁵, whereas alkene hydroaminations by lanthanide catalysts have been largely limited to reactions occurring by intramolecular insertions^{90,91}. Because lanthanide-catalysed olefin hydroamination is slower than lanthanide-catalysed olefin polymerization, the rate of olefin insertions into lanthanide amides seems to be slower than that of olefin insertions into lanthanide alkyls. This slower rate of insertion into lanthanide amides can be understood by the reaction's requirement to break an M–N bond that is stronger than the M–C bond. The M–N bond is stronger in this case because of a favourable match of a hard metal with a hard ligand and donation of the non-bonded electrons from the amido group into unoccupied orbitals on the metal.

If the relative rates of insertions of olefins into lanthanide–amide and lanthanide–alkyl complexes are affected by the hardness of the lanthanide metal and the ability of this metal to accept electron density from the electron pair on nitrogen, then the relative rates for insertions of olefins into amide and alkyl complexes of softer, more electron-rich late transition metals are likely to be different. Although insertions of olefins into cationic or electrophilic neutral palladium–alkyl and nickel–alkyl complexes can be fast^{96,97}, two recent studies suggest that the insertions of olefins into more electron-rich, late-transition-metal amides and alkoxides can be faster than into the analogous alkyl complexes.

Certain catalytic amidoarylations and alkoxyarylations of olefins reveal the relative rates of insertion into M–amido and M–aryl linkages^{98,99}. These reactions occur through arylpalladium amido and arylpalladium alkoxo complexes, and in each case the product from insertion of the olefin into the M–amido or M–alkoxo bond is formed. In one particularly revealing experiment (Fig. 9a), the selectivity from a catalytic alkoxyarylation process involving a palladium centre bound by one alkoxo and

one aryl ligand showed that insertion into the palladium–alkoxo bond was faster than insertion into the palladium–aryl bond⁹⁹.

A second set of data (Fig. 9b) reveals the relative rates of olefin insertions into discrete rhodium–alkoxo and rhodium–amido complexes^{31,32}. In one study, intermolecular insertion of an olefin into a rhodium–amido complex forms a rhodium hydride and an enamine as the final products after β -hydrogen elimination from the initially formed aminoalkyl species³¹. In a closely related study, intramolecular insertion of a coordinated alkene into a rhodium alkoxide forms a cyclic vinyl ether after β -hydrogen elimination from the resultant alkoxyalkyl intermediate³². The corresponding rhodium methyl complex does not insert alkenes. Thus, the rates for insertions of olefins into these rhodium alkoxides and amides seem to be faster than for insertions into the analogous alkyls, and these early data imply that this trend applies to the reactions of many other late-transition-metal–alkyl, –alkoxo and –amido complexes. These rates could be controlled by the electron pair of the ligand as summarized in Fig. 9c. First, the M–heteroatom bonds in the starting, low-valent alkoxide or amide would probably be destabilized, relative to the analogous alkyl complexes, by repulsion between the filled metal *d* orbital and the filled amide or alkoxide *p* orbital. Second, the M–heteroatom bond need not be fully cleaved during insertions into metal alkoxides and amides because the insertions can form metallocyclic products in which the covalent M–Y bond of the starting amido or alkoxo complex has been converted into an M–Y dative bond (where Y is NH or O). Because the M–N and M–O bonds have not been fully cleaved in this process, the barrier for insertion can be envisaged to be lower than for insertion into the related alkyl complex. Thus, several differences in the properties of organometallic and hetero-organometallic complexes can make amido and alkoxo complexes more reactive than M–alkyl complexes towards insertions of alkenes.

Conclusions

Catalytic reactions that are widely used in organic synthesis occur with a broad scope of substrates, a high tolerance of the reaction for auxiliary functional groups, high turnover numbers, fast rates, inexpensive reagents and minimal side products. Although great progress has been made in the coupling of amines with aryl halides, the additions of amines to olefins, the oxidation of olefins and the functionalization of C–H bonds, further progress is needed before all of these reactions will become everyday tools of the synthetic chemist.

The additions of amines to aryl halides occur with sufficient scope and efficiency that it is now a commonly used synthetic method in both academic and industrial settings. However, advances in the other catalytic reactions discussed in this review are needed for them to become a common synthetic tool. For example, coupling of aryl halides with weaker heteroatom nucleophiles, such as alcohols, does not yet occur with an acceptable scope and efficiency. Likewise, intramolecular hydroamination of alkenes needs to occur with more hindered alkenes, with control of stereochemistry and with much lower amounts of catalyst, and it needs to be developed into an intermolecular process. The functionalization of C–H bonds needs to be developed into a process that occurs with milder and less expensive reagents, and it needs to occur with much lower amounts of catalyst and additives. It is anticipated that an increased understanding of the factors that control the elementary reactions of these catalytic cycles will greatly accelerate the discovery of new catalysts that overcome these current limitations.

The establishment of these principles and their use in generating improved reactions will bring the field of homogeneous catalysis closer to a stage at which new catalysts and catalytic reactions can be designed. Although a framework has been established to plan synthetic sequences¹⁰⁰, a similar conceptual framework to design catalysts and catalytic reactions is far from established. As additional reactions to form C–C and C–heteroatom bonds are discovered and further refined, and as our understanding of the factors that control the rates and selectivities of these processes improves, the field of homogeneous catalysis will move closer to the point at which researchers can make rational choices about how to modify existing catalysts and can begin to design new catalysts for discovering the next set of dream reactions. ■

- Muci, A. R. & Buchwald, S. L. Practical palladium catalysts for C–N and C–O bond formation. *Top. Curr. Chem.* **219**, 131–209 (2002).
This paper reviews the primary literature on cross-coupling to form C–N and C–O bonds.
- Hartwig, J. F. in *Handbook of Organopalladium Chemistry for Organic Synthesis* Vol. 1 (ed. Negishi, E. I.) 1051–1096 (Wiley-Interscience, 2002).
- Hartwig, J. F. in *Handbook of Organopalladium Chemistry for Organic Synthesis* Vol. 1 (ed. Negishi, E. I.) 1097–1106 (Wiley-Interscience, 2002).
- Dick, A. R. & Sanford, M. S. Transition metal catalyzed oxidative functionalization of carbon–hydrogen bonds. *Tetrahedron* **62**, 2439–2463 (2006).
This paper reviews various modern approaches to C–H bond functionalization to form C–heteroatom bonds.
- Hull, K. L., Anani, W. Q. & Sanford, M. S. Palladium-catalyzed fluorination of carbon–hydrogen bonds. *J. Am. Chem. Soc.* **128**, 7134–7135 (2006).
- Giri, R., Chen, X. & Yu, J. Q. Palladium-catalyzed asymmetric iodination of unactivated C–H bonds under mild conditions. *Angew. Chem. Int. Edn Engl.* **44**, 2112–2115 (2005).
- Chen, H., Schlecht, S., Semple, T. C. & Hartwig, J. F. Thermal, catalytic, regioselective functionalization of alkanes. *Science* **287**, 1995–1997 (2000).
This paper reports the only catalytic process to functionalize terminal alkyl C–H bonds selectively.
- Jira, R. in *Applied Homogeneous Catalysis with Organometallic Compounds: A Comprehensive Handbook in Two Volumes* (eds Cornils, B. & Herrmann, W. A.) 386–405 (Wiley-VCH, 2002).
- Liu, G. & Stahl, S. S. Two-faced reactivity of alkenes: *cis*- versus *trans*-aminopalladation in aerobic Pd-catalyzed intramolecular aza-Wacker reactions. *J. Am. Chem. Soc.* **129**, 6328–6335 (2007).
- Muller, T. E. in *Encyclopedia of Catalysis* Vol. 3 (ed. Horváth, I. T.) 518–541 (Wiley-Interscience, 2003).
- King, A. O. & Yasuda, N. Palladium-catalyzed cross-coupling reactions in the synthesis of pharmaceuticals. *Topics Organomet. Chem.* **6**, 205–245 (2004).
- de Vries, J. G. The Heck reaction in the production of fine chemicals. *Can. J. Chem.* **79**, 1086–1092 (2001).
- Aizawa, M., Yamada, T., Shinohara, H., Akagi, K. & Shirakawa, H. Electrochemical fabrication of a polypyrrole–polythiophene p–n-junction diode. *J. Chem. Soc., Chem. Commun.* **17**, 1315–1317 (1986).
- Jen, K.-Y., Miller, G. G. & Elsenbaumer, R. L. Highly conducting, soluble, and environmentally-stable poly(3-alkylthiophenes). *J. Chem. Soc., Chem. Commun.* **17**, 1346–1347 (1986).
- Sato, M.-A., Tanaka, S. & Kaeriyama, K. Soluble conducting polythiophenes. *J. Chem. Soc., Chem. Commun.* **11**, 873–874 (1986).
- McQuade, D. T., Pullen, A. E. & Swager, T. M. Conjugated polymer-based chemical sensors. *Chemical Reviews* **100**, 2537–2574 (2000).
- Hasenhuettl, G. L. in *Kirk-Othmer Encyclopedia of Chemical Technology*, doi:10.1002/0471238961.0601201908011905.a01.pub2 (Wiley, 2005).
- Blaser, H. U. & Schmidt, E. *Asymmetric Catalysis on Industrial Scale: Challenges, Approaches and Solutions*. (Wiley-VCH, 2004).
- Hoveyda, A. H. & Zhugralin, A. R. The remarkable metal-catalysed olefin metathesis reaction. *Nature* **450**, 243–251 (2007).
- Kakiuchi, F. & Chatani, N. Catalytic methods for C–H bond functionalization: application in organic synthesis. *Adv. Synth. Catal.* **345**, 1077–1101 (2003).
- Labinger, J. A. & Bercaw, J. E. Understanding and exploiting C–H bond activation. *Nature* **417**, 507–514 (2002).
- Goldman, A. S. *et al.* Catalytic alkane metathesis by tandem alkane dehydrogenation olefin metathesis. *Science* **312**, 257–261 (2006).
- Crabtree, R. H. *The Organometallic Chemistry of the Transition Metals* 4th edn (Wiley, 2005).
- Collman, J. P., Hegedus, L. S., Norton, J. R. & Finke, R. G. *Principles and Applications of Organotransition Metal Chemistry* (University Science Books, 1987).
- Ojima, I., Li, Z. & Zhu, J. in *The Chemistry Of Organic Silicon Compounds* Vol. 2 (eds Rappaport, Z. & Apeloig, Y.) 1687–1792 (Wiley, 1998).
- Hartwig, J. F. Carbon–heteroatom bond-forming reductive elimination of amines, ethers, and sulfides. *Acc. Chem. Res.* **31**, 852–860 (1998).
- Hartwig, J. F. Electronic effects on reductive elimination to form carbon–carbon and carbon–heteroatom bonds from palladium(II) complexes. *Inorg. Chem.* **46**, 1936–1947 (2007).
This paper reviews how the properties of a heteroatom affect the rates of reductive elimination to form C–heteroatom bonds.
- Hillhouse, G. L. & Bercaw, J. E. Reactions of water and ammonia with bis(pentamethylcyclopentadienyl) complexes of zirconium and hafnium. *J. Am. Chem. Soc.* **106**, 5472–5478 (1984).
- Zhao, J., Goldman, A. S. & Hartwig, J. F. Oxidative addition of ammonia to form a stable monomeric amido hydride complex. *Science* **307**, 1080–1082 (2005).
- Casalnuovo, A. L., Calabrese, J. C. & Milstein, D. Rational design in homogeneous catalysis. Ir(I)-catalyzed addition of aniline to norbornylene via N–H activation. *J. Am. Chem. Soc.* **110**, 6738–6744 (1988).
- Zhao, P. J., Krug, C. & Hartwig, J. F. Transfer of amido groups from isolated rhodium(I) amides to alkenes and vinylarenes. *J. Am. Chem. Soc.* **127**, 12066–12073 (2005).
- Zhao, P., Incarvito, C. D. & Hartwig, J. F. Carbon–oxygen bond formation between a terminal alkoxo ligand and a coordinated olefin. Evidence for olefin insertion into a rhodium alkoxide. *J. Am. Chem. Soc.* **128**, 9642–9643 (2006).
- Polse, J. L., Andersen, R. A. & Bergmann, R. G. Reactivity of a terminal Ti(IV) imido complex toward alkenes and alkynes: cycloaddition vs C–H activation. *J. Am. Chem. Soc.* **120**, 13405–13414 (1998).
- Fulton, J. R., Holland, A. W., Fox, D. J. & Bergman, R. G. Formation, reactivity, and properties of nondeative late transition metal–oxygen and –nitrogen bonds. *Acc. Chem. Res.* **35**, 44–56 (2002).
- Bryndza, H. & Tam, W. Monomeric metal hydroxides, alkoxides, and amides of the late transition metals: synthesis, reactions, and thermochemistry. *Chem. Rev.* **88**, 1163–1188 (1988).
- Hartwig, J. F. in *Modern Arene Chemistry* (ed. Astruc, C.) 107–168 (Wiley-VCH, 2002).
- Boger, D. L. & Panek, J. S. Palladium (0) mediated β -carboline synthesis: preparation of the CDE ring system of lavendamyacin. *Tetrahedron Lett.* **25**, 3175–3178 (1984).
- Kosugi, M., Kameyama, M. & Migita, T. Palladium-catalyzed aromatic amination of aryl bromides with *N,N*-diethylamino-tributyltin. *Chem. Lett.* **12**, 927–928 (1983).
- Guram, A. S., Rennels, R. A. & Buchwald, S. L. A simple catalytic method for the conversion of aryl bromides to arylamines. *Angew. Chem. Int. Edn Engl.* **34**, 1348–1350 (1995).
- Louie, J. & Hartwig, J. F. Palladium-catalyzed synthesis of arylamines from aryl halides. mechanistic studies lead to coupling in the absence of tin reagents. *Tetrahedron Lett.* **36**, 3609–3612 (1995).
- Wolfe, J. P., Wagaw, S. & Buchwald, S. L. An improved catalyst system for aromatic carbon–nitrogen bond formation: the possible involvement of bis(phosphine) palladium complexes as key intermediates. *J. Am. Chem. Soc.* **118**, 7215–7216 (1996).
- Driver, M. S. & Hartwig, J. F. A second generation catalyst for aryl halide amination: mixed secondary amines from aryl halides and primary amines catalyzed by (DPPF)PdCl₂. *J. Am. Chem. Soc.* **118**, 7217–7218 (1996).
- Nishiyama, M., Yamamoto, T. & Koie, Y. Synthesis of *N*-arylpiperazines from aryl halides and piperazine under a palladium tri-*tert*-butylphosphine catalyst. *Tetrahedron Lett.* **39**, 617–620 (1998).
- Hartwig, J. F., Kawatsura, M., Hauck, S. I., Shaughnessy, K. H. & Alcazar-Roman, L. M. Room-temperature palladium-catalyzed amination of aryl bromides and chlorides and extended scope of aromatic C–N bond formation with a commercial ligand. *J. Org. Chem.* **64**, 5575–5580 (1999).
- Stambuli, J. P., Kuwano, R. & Hartwig, J. F. Unparalleled rates for the activation of aryl chlorides. Coupling with amines and boronic acids in minutes at room temperature. *Angew. Chem. Int. Edn Engl.* **41**, 4746–4747 (2002).
- Zapf, A. *et al.* Practical synthesis of new and highly efficient ligands for the Suzuki reaction of aryl chlorides. *Chem. Commun.* **10**, 38–39 (2004).
- Singer, R. A., Dore, M. L., Sieser, J. E. & Berliner, M. A. Development of nonproprietary phosphine ligands for the Pd-catalyzed amination reaction. *Tetrahedron Lett.* **47**, 3727–3731 (2006).
- Kataoka, N., Shelby, Q., Stambuli, J. P. & Hartwig, J. F. Air stable, sterically hindered ferrocenyl dialkylphosphines for palladium-catalyzed C–C, C–N and C–O bond-forming cross-couplings. *J. Org. Chem.* **67**, 5553–5566 (2002).
- Shen, Q., Shekhar, S., Stambuli, J. P. & Hartwig, J. F. Highly reactive, general, and long-lived catalysts for coupling heteroaryl and aryl chlorides with primary nitrogen nucleophiles. *Angew. Chem. Int. Edn Engl.* **44**, 1371–1375 (2004).
- Shen, Q. & Hartwig, J. F. Palladium-catalyzed coupling of ammonia and lithium amide. *J. Am. Chem. Soc.* **128**, 10028–10029 (2006).
- Driver, M. S. & Hartwig, J. F. Carbon–nitrogen bond-forming reductive elimination of arylamines from Pd(II). *J. Am. Chem. Soc.* **119**, 8232–8245 (1997).
- Fujita, K. I. *et al.* Organometallic chemistry from amidate complexes. Reductive elimination of *N*-aryl amidates from palladium(II). *J. Am. Chem. Soc.* **128**, 9044–9045 (2006).
- Mann, G., Incarvito, C., Rheingold, A. L. & Hartwig, J. F. Palladium-catalyzed C–O coupling involving unactivated aryl halides. Sterically induced reductive elimination to form the C–O bond in diaryl ethers. *J. Am. Chem. Soc.* **121**, 3224–3225 (1999).
- Stambuli, J. P., Weng, Z. Q., Incarvito, C. D. & Hartwig, J. F. Reductive elimination of ether from T-shaped, monomeric arylpalladium alkoxides. *Angew. Chem. Int. Edn Engl.* **46**, 7674–7677 (2007).

55. Mann, G., Barañano, D., Hartwig, J. F., Rheingold, A. L. & Guzei, I. A. Carbon-sulfur bond-forming reductive elimination involving sp -, sp^2 - and sp^3 -hybridized carbon. Mechanism, steric effects, and electronic effects on sulfide formation. *J. Am. Chem. Soc.* **120**, 9205–9219 (1998).
56. Holland, P. L., Andersen, R. A. & Bergman, R. G. Application of the E-C approach to understanding the bond energies thermodynamics of late-metal amido, aryloxo and alkoxo complexes: an alternative to $p\pi/d\pi$ repulsion. *Comments Inorg. Chem.* **21**, 115–129 (1999).
57. Culkin, D. A. & Hartwig, J. F. Carbon-carbon bond-forming reductive elimination from arylpalladium complexes containing functionalized alkyl groups. Influence of ligand steric and electronic properties on structure, stability, and reactivity. *Organometallics* **23**, 3398–3416 (2004).
58. Huheey, J. E., Keiter, E. A. & Keiter, R. L. *Inorganic Chemistry* 4th edn (Harper Collins, 1993).
59. Williams, B. S. & Goldberg, K. I. Studies of reductive elimination reactions to form carbon-oxygen bonds from Pt(IV) complexes. *J. Am. Chem. Soc.* **123**, 2576–2587 (2001).
60. Pawlikowski, A. V., Getty, A. D. & Goldberg, K. I. Alkyl carbon-nitrogen reductive elimination from platinum(IV)-sulfonamide complexes. *J. Am. Chem. Soc.* **129**, 10382–10393 (2007).
61. Dick, A. R., Kampf, J. W. & Sanford, M. S. Unusually stable palladium(IV) complexes: detailed mechanistic investigation of C–O bond-forming reductive elimination. *J. Am. Chem. Soc.* **127**, 12790–12791 (2005).
62. Bistri, O., Correa, A. & Bolm, C. Iron-catalyzed C–O cross-couplings of phenols with aryl iodides. *Angew. Chem. Int. Edn Engl.* **47**, 586–588 (2008).
63. Correa, A. & Bolm, C. Iron-catalyzed *N*-arylation of nitrogen nucleophiles. *Angew. Chem. Int. Edn Engl.* **46**, 8862–8865 (2007).
64. Beletskaya, I. P. & Cheprakov, A. V. Copper in cross-coupling reactions — the post-Ullmann chemistry. *Coord. Chem. Rev.* **248**, 2337–2364 (2004).
65. Shilov, A. E. & Shul'pin, G. B. Activation of C–H bonds by metal complexes. *Chem. Rev.* **97**, 2879–2932 (1997).
66. Periana, R. A. *et al.* Platinum catalysts for the high-yield oxidation of methane to a methanol derivative. *Science* **280**, 560–564 (1998).
67. Sen, A. Catalytic functionalization of carbon-hydrogen and carbon-carbon bonds in protic media. *Acc. Chem. Res.* **31**, 550–557 (1998).
68. Stahl, S. S., Labinger, J. A. & Bercaw, J. E. Homogeneous oxidation of alkanes by electrophilic late transition metals. *Angew. Chem. Int. Edn Engl.* **37**, 2181–2192 (1998).
69. Groves, J. T. High-valent iron in chemical and biological oxidations. *J. Inorg. Biochem.* **100**, 434–447 (2006).
70. Wenzel, T. T. & Bergman, R. G. Inter- and intramolecular insertion of rhenium into carbon-hydrogen bonds. *J. Am. Chem. Soc.* **108**, 4856–4867 (1986).
71. Jones, W. D. & Feher, F. J. The mechanism and thermodynamics of alkane and arene carbon-hydrogen bond activation in $(C_5Me_5)Rh(PMe_3)(R)H$. *J. Am. Chem. Soc.* **106**, 1650–1663 (1985).
72. Arndtsen, B. A., Bergman, R. G., Mobley, T. A. & Peterson, T. H. Selective intermolecular carbon-hydrogen bond activation by synthetic metal complexes in homogeneous solution. *Acc. Chem. Res.* **28**, 154 (1995).
73. Whitfield, S. R. & Sanford, M. S. Reactivity of Pd(II) complexes with electrophilic chlorinating reagents: isolation of Pd(IV) products and observation of C–Cl bond-forming reductive elimination. *J. Am. Chem. Soc.* **129**, 15142–15143 (2007).
74. Giri, R., Chen, X. & Yu, J. Q. Palladium-catalyzed asymmetric iodination of unactivated C–H bonds under mild conditions. *Angew. Chem. Int. Edn Engl.* **44**, 2112–2115 (2005).
75. Desai, L. V., Hull, K. L. & Sanford, M. S. Palladium-catalyzed oxygenation of unactivated sp^3 C–H bonds. *J. Am. Chem. Soc.* **126**, 9542–9543 (2004).
76. Davies, D. L., Donald, S. M. A. & Macgregor, S. A. Computational study of the mechanism of cyclometalation by palladium acetate. *J. Am. Chem. Soc.* **127**, 13754–13755 (2005).
77. Ishiyama, T. & Miyaura, N. Transition metal-catalyzed borylation of alkanes and arenes via C–H activation. *J. Organomet. Chem.* **680**, 3–11 (2003).
78. Hartwig, J. F. *et al.* Rhodium-boryl complexes in the catalytic, terminal functionalization of alkanes. *J. Am. Chem. Soc.* **127**, 2538–2552 (2005).
79. Henry, P. M. in *Handbook of Organopalladium Chemistry for Organic Synthesis* (ed. Negishi, E. I.) (Wiley-Interscience, 2002).
80. Francis, J. W. & Henry, P. M. Palladium(II)-catalyzed exchange and isomerization reactions. 14. Kinetics and stereochemistry of the isomerization and water exchange of 2-(methyl- d_3)-4-methyl-1,1,1,5,5,5-hexafluoro-3-penten-2-ol in aqueous solution catalyzed by $PdCl_4^{2-}$: two new mechanistic probes for catalytic chemistry. *Organometallics* **10**, 3498–3503 (1991).
81. Francis, J. W. & Henry, P. M. Palladium(II)-catalyzed exchange and isomerization reactions. 15. Kinetics and stereochemistry of the isomerization and water exchange of 2-(methyl- d_3)-4-methyl-1,1,1,5,5,5-hexafluoro-3-penten-2-ol in aqueous solution catalyzed by $PdCl_4^{2-}$ at high chloride concentrations. *Organometallics* **11**, 2832–2836 (1992).
82. Hamed, O., Thompson, C. & Henry, P. M. Stereochemistry of the Wacker reaction: modes of addition of hydroxide, methoxide, and phenyl at high and low Cl^- . A study using chirality transfer. *J. Org. Chem.* **62**, 7082–7083 (1997).
83. Hayashi, T., Yamasaki, K., Mimura, M. & Uozumi, Y. Deuterium-labeling studies establishing stereochemistry at the oxypalladation step in Wacker-type oxidative cyclization of an *O*-allylphenol. *J. Am. Chem. Soc.* **126**, 3036–3037 (2004).
84. Trend, R. M., Ramtohl, Y. K. & Stoltz, B. M. Oxidative cyclizations in a nonpolar solvent using molecular oxygen and studies on the stereochemistry of oxypalladation. *J. Am. Chem. Soc.* **127**, 17778–17788 (2005).
85. Hong, S. & Marks, T. J. Organolanthanide-catalyzed hydroamination. *Acc. Chem. Res.* **37**, 673–686 (2004).
- This paper reviews the development of some of the most active catalysts for alkene hydroamination, which function by insertions of alkenes into M–amide bonds.**
86. Amin, S. B. & Marks, T. J. Organolanthanide-catalyzed synthesis of amine-capped polyethylenes. *J. Am. Chem. Soc.* **129**, 10102–10103 (2007).
87. Roesky, P. W. & Muller, T. E. Enantioselective catalytic hydroamination of alkenes. *Angew. Chem. Int. Edn Engl.* **42**, 2708–2710 (2003).
88. Hultzschi, K. C. Transition metal-catalyzed asymmetric hydroamination of alkenes (AHA). *Adv. Synth. Catal.* **347**, 367–391 (2005).
89. Kim, J. Y. & Livinghouse, T. Enantioselective intramolecular alkene hydroaminations catalyzed by yttrium complexes of axially chiral bis(thiolate) ligands. *Org. Lett.* **7**, 1737–1739 (2005).
90. Gagné, M. R. & Marks, T. J. Organolanthanide-catalyzed hydroamination. Facile regioselective cyclization of unprotected amino olefins. *J. Am. Chem. Soc.* **111**, 4108–4109 (1989).
91. Gagne, M. R., Stern, C. L. & Marks, T. J. Organolanthanide-catalyzed hydroamination — a kinetic, mechanistic, and diastereoselectivity study of the cyclization of *N*-unprotected amino olefins. *J. Am. Chem. Soc.* **114**, 275–294 (1992).
92. Walsh, P. J., Baranger, A. M. & Bergman, R. G. Stoichiometric and catalytic hydroamination of alkynes and allene by zirconium bisamides $Cp_2Zr(NHR)_2$. *J. Am. Chem. Soc.* **114**, 1708–1719 (1992).
93. Wood, M. C., Leitch, D. C., Yeung, C. S., Kozak, J. A. & Schafer, L. L. Chiral neutral zirconium amide complexes for the asymmetric hydroamination of alkenes. *Angew. Chem. Int. Edn Engl.* **46**, 354–358 (2007).
94. Gott, A. L., Clarke, A. J., Clarkson, G. J. & Scott, P. Catalytic alkene cyclohydroamination via an imido mechanism. *Chem. Commun.* **12**, 1422–1424 (2008).
95. Jeske, G. *et al.* Highly reactive organolanthanides — systematic routes to and olefin chemistry of early and late bis(pentamethylcyclopentadienyl) 4f hydrocarbyl and hydride complexes. *J. Am. Chem. Soc.* **107**, 8091–8103 (1985).
96. Ittel, S. D., Johnson, L. K. & Brookhart, M. Late-metal catalysts for ethylene homo- and copolymerization. *Chem. Rev.* **100**, 1169–1203 (2000).
97. Mecking, S. Olefin polymerization by late transition metal complexes — a root of Ziegler catalysts gains new ground. *Angew. Chem. Int. Edn Engl.* **40**, 534–540 (2001).
98. Wolfe, J. P. Palladium-catalyzed carboetherification and carboamination reactions of γ -hydroxy- and γ -aminoalkenes for the synthesis of tetrahydrofurans and pyrrolidines. *Eur. J. Org. Chem.* **4**, 571–582 (2007).
99. Nakhla, J. S., Kampf, J. W. & Wolfe, J. P. Intramolecular Pd-catalyzed carboetherification and carboamination. Influence of catalyst structure on reaction mechanism and product stereochemistry. *J. Am. Chem. Soc.* **128**, 2893–2901 (2006).
100. Corey, E. J. & Cheng, X.-M. *The Logic of Chemical Synthesis* (Wiley-Interscience, 1995).

Acknowledgements I thank the National Institutes of Health, the US Department of Energy and the National Science Foundation for funding my research related to the catalytic formation of carbon-heteroatom bonds. I also thank my co-workers who helped to formulate the concepts included in the Review, and E. Alexanian for suggestions and thorough editing of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to the author (jhartwig@uiuc.edu).

Natural products as inspiration for the development of asymmetric catalysis

Justin T. Mohr¹, Michael R. Krout¹ & Brian M. Stoltz¹

Biologically active natural products often contain particularly challenging structural features and functionalities in terms of synthesis. Perhaps the greatest difficulties are those caused by issues of stereochemistry. A useful strategy for synthesizing such molecules is to devise methods of bond formation that provide opportunities for using enantioselective catalysis. In using this tactic, the desire for a particular target structure ultimately drives the development of catalytic methods. New enantioselective catalytic methods contribute to a greater fundamental understanding of how bonds can be constructed and lead to valuable synthetic technologies that are useful for a variety of applications.

The lack of methods available for installing functionalities or structural motifs during chemical synthesis can at first be frustrating. However, retrosynthetic analysis¹, a way of viewing the target molecule as a series of structurally simpler precursors, can greatly aid in planning how to generate a valuable chemical substance. Despite this, difficulties in preparing materials enriched in a particular enantiomer persist because of the limited number of catalytic enantioselective transformations available^{2–6}. One fruitful strategy is to design a synthesis that depends on a bond-forming reaction for which there is no known enantioselective variant. This approach thus provides the impetus for developing novel transformations and leads to a greater understanding of methods of bond construction and catalysis.

In this review, we describe several recent examples of novel catalytic enantioselective transformations that illustrate the effectiveness of this strategy for preparing important structural motifs found in biologically active molecules. Each of these transformations has contributed not only an effective means of generating a particular target structure but also a useful new tool for a variety of applications in synthetic chemistry.

Historical overview of enantioselective methods

To provide an overview of established catalytic enantioselective methods that have been developed for total synthesis, several notable examples of enantioselective reactions in total synthesis are highlighted in Fig. 1. In each of these cases, the target molecules posed particular challenges that had yet to be solved by enantioselective catalysis. Although, in some instances (for example, the Diels–Alder reaction, Fig. 1a), the methods were developed before their first application in total synthesis, the demonstrated value of the transformation highlighted the need for enantioselective variants. Following the development of the [4 + 2] cycloaddition reaction⁷ in the 1920s, studies of this transformation elucidated several key facets of the stereochemical outcome of the reaction (for example, the ‘*endo*’ rule, regioselectivity and diastereoselectivity). These intrinsic stereochemical control elements proved useful when the Diels–Alder reaction was first featured in a total synthesis with Gilbert Stork’s stereocontrolled synthesis of cantharidin⁸ in 1951. Subsequently, the thermal Diels–Alder reaction

was used for several total syntheses, perhaps most famously in Robert Woodward’s landmark synthesis of reserpine⁹. Enantioselectivity in this transformation remained elusive, however, and perhaps was considered unattainable at the time.

One key practical improvement in the Diels–Alder reaction was the discovery that Lewis acids markedly increased the reaction rate¹⁰. Many laboratories sought to exploit this and to develop asymmetric versions of the Diels–Alder reaction catalysed by chiral Lewis acids, culminating in a report of the first highly enantioselective catalytic Diels–Alder reaction¹¹ in 1979. The interface between reaction development, study of the mechanism, and synthesis is readily apparent from the multitude of chiral Diels–Alder catalysts and accompanying enantioselective total syntheses that have been reported (see refs 12–14 for examples). These successes validate the extensive efforts directed at realizing this important goal.

Other methods were developed to address more general problems in synthesis (for example, synthesis of chiral alcohols by means of enantioselective ketone reduction, Fig. 1a); however, the key structures are embedded in a variety of important natural products and pharmaceutical compounds. In the case of Elias J. Corey’s approach to the synthesis of prostaglandins¹⁵, first reported in the 1960s, control of the configuration of the sidechain allylic alcohol at C15 required stoichiometric, chiral reducing agents until a solution to this long-standing problem was found in the 1980s (ref. 16). Interestingly, the oxazaborolidine catalyst discovered in these explorations has had other varied applications in synthesis and catalysis^{13,17}, demonstrating the versatility of privileged molecular frameworks¹⁸ for enantioselective catalysis.

The practical application of enantioselective catalysis is apparent in myriad industrial applications (for example, Fig. 1b–d), for which the limits of catalysis must be examined to minimize costs. Important industrial applications include the synthesis of chiral building blocks (for example, amino acids¹⁹, Fig. 1b), novel biologically active pharmaceuticals (for example, Crixivan (indinavir sulphate)²⁰, Fig. 1c) and commodity chemicals (cheap chemicals sold in bulk) with various important uses (for example, menthol²¹, Fig. 1d). Only the most efficient methods are feasible for large-scale industrial synthesis, and in many ways these protocols represent the pinnacle of modern

¹The Arnold and Mabel Beckman Laboratories of Chemical Synthesis, Division of Chemistry and Chemical Engineering, California Institute of Technology, 1200 East California Boulevard, MC 164-30, Pasadena, California 91125, USA.

enantioselective catalysis^{22–24}. A viable commercial operation must account for more than simply effective asymmetric induction; factors including turnover frequency, catalyst availability, catalyst recovery, catalyst toxicity and feasible large-scale handling procedures must all be considered for industrial applications. These daunting challenges underscore the demand for increasingly efficient catalyst systems.

To maximize the usefulness of the stereochemistry attained by these key asymmetric transformations, subsequent diastereoselective reactions may be used to control the formation of many stereocentres based on a single enantioselective transformation (for example, Fig. 1e). The Hajos–Parrish ketone, first prepared in the context of steroid synthesis, has been used extensively in other synthetic efforts and has proved to be a versatile chiral-pool starting material^{25–27}. The amino-acid catalyst system developed for this intramolecular aldol condensation provided a sound basis for the recent use of organic molecules as catalysts for a variety of enantioselective transformations (see the subsection ‘Flustramine B’).

The use of several different enantioselective reactions to prepare enantio-enriched fragments of complex molecules improves efficiency through convergency. The importance of this strategy is shown by the variety of extraordinarily complex polyketide natural products that have been prepared through asymmetric intermolecular aldol reactions (for example, phorboxazole B²⁸, Fig. 1f). The challenging structure of these molecules has required the development of several related protocols to address the subtle differences in substitution patterns and functionality present in substrates, and, despite many successes, studies are ongoing²⁹.

Recent developments in enantioselective catalysis

In this section, we describe recent representative developments made by using this approach — that is, by using target structures to inspire the development of enantioselective catalysts — for the construction of biologically important target molecules. Most of these methods involve the formation of a carbon–carbon bond, the fundamental structure of organic molecules. These cases were selected to illustrate some of the latest developments in enantioselective catalysis for complex molecule synthesis. Special attention has been given to reactions that address some of the most important challenges in synthetic chemistry today: increasing functional-group tolerance, generating new carbocyclic and heterocyclic rings, and forming all-carbon quaternary stereocentres. The examples are also intended to show the important symbiosis between total synthesis and method development, and to show that improvements in one branch of synthetic chemistry have an impact on the others.

Januvia

Catalytic enantioselective hydrogenation has become one of the most effective and powerful methods for the synthesis of chiral α -amino acids for numerous applications¹⁹. Over the past decade, the usefulness of the homologous building blocks, β -amino acids, in pharmaceutical, agrochemical, β -peptide and natural substances has become evident, highlighting the need for a general and effective means for their preparation^{30,31}. Undoubtedly, the implementation of a catalytic asymmetric

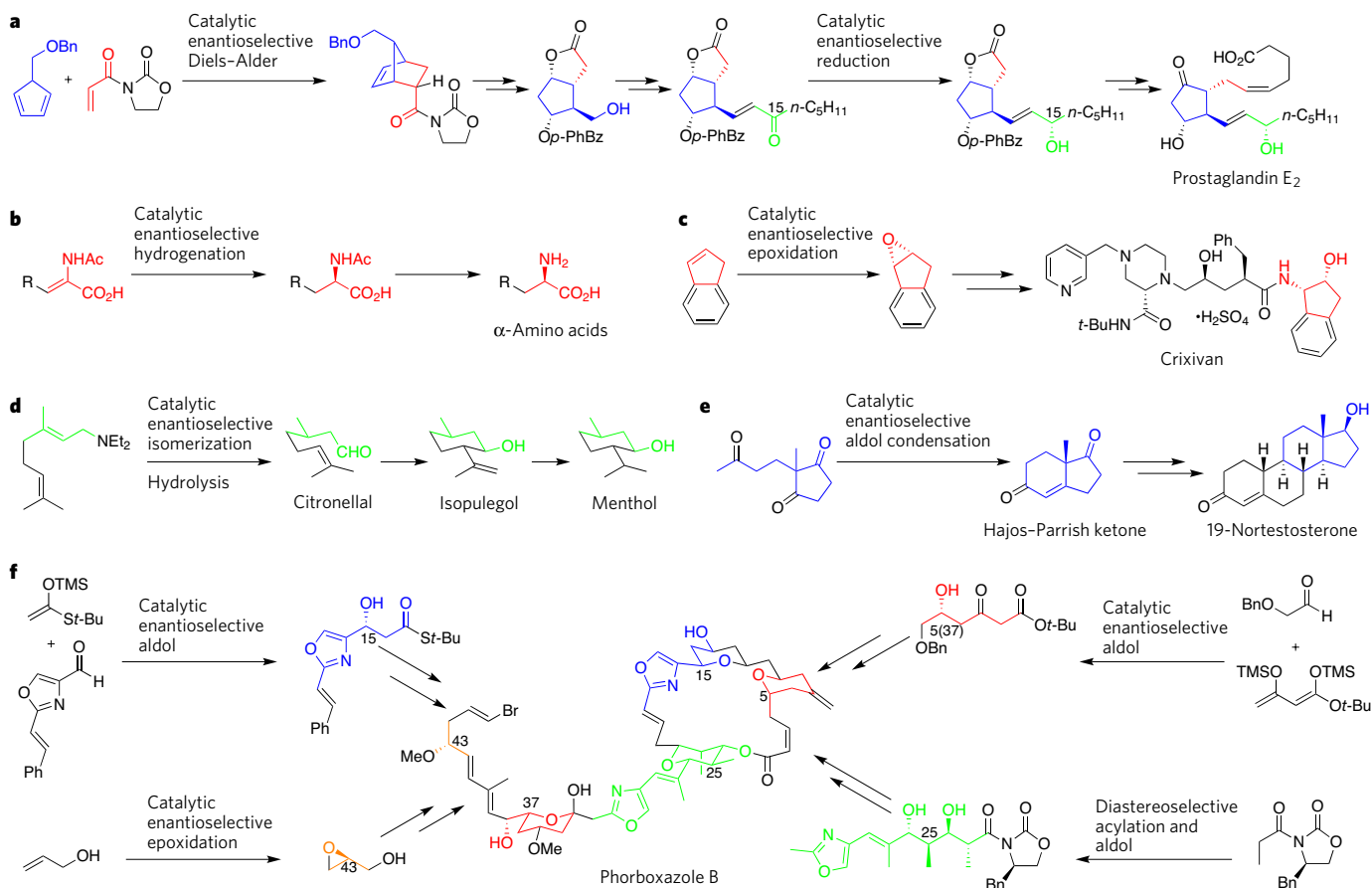


Figure 1 | Selected examples of enantioselective catalysis in total synthesis. Natural products and pharmaceuticals often have structural features and functionalities that are challenging to synthesize. This particularly applies to stereochemical elements. Selected examples of successful enantioselective syntheses are illustrated. **a**, Corey's synthesis of prostaglandin E₂ utilizing an asymmetric Diels–Alder reaction¹³ and carbonyl reduction¹⁶. Various industrial applications have explored the limits of catalysis and efficiency to minimize cost. These include asymmetric hydrogenation towards α -amino-acid building blocks¹⁹ (**b**), asymmetric epoxidation towards Merck's HIV

protease inhibitor Crivivan²⁰ (**c**), and asymmetric isomerization of allyl amines en route to commodity chemicals²¹ (**d**). The use of stereochemical information in valuable intermediates can be maximized through convergency, as illustrated by the asymmetric construction of C–C bonds in the amino-acid-catalysed intramolecular aldol condensation towards important steroid building blocks²⁷ (**e**) and in various enantioselective intermolecular aldol reactions towards fragments of the cytostatic agent phorboxazole B²⁸ (**f**). Ac, acetyl; Bn, benzyl; Bz, benzoyl; Et, ethyl; Me, methyl; Ph, phenyl; *t*-Bu, *tert*-butyl; TMS, trimethylsilyl.

hydrogenation of *N*-acyl- β -enamino esters seemed to be the most efficient pathway towards their synthesis, although initial investigations achieved poor selectivities³². Additional syntheses using the chiral pool, auxiliaries and more recently the catalytic asymmetric generation of C–C and C–N bonds have been successful in satisfying the increased demand for β -amino acids³¹. These valuable methods allow flexible strategies for the synthesis of a variety of analogues; however, most examples are limited by the requirement for further chemical manipulation that is often necessary to produce the functionality of the desired β -amino acids.

Despite initial difficulties, the asymmetric hydrogenation of *N*-acyl- β -enamino esters has been developed into a useful method over the past 15 years^{33,34}. This fruitful endeavour has demonstrated that several transition metal and ligand combinations are competent for preparing *N*-acyl- β -amino acids with good-to-excellent enantioselectivities. A notable drawback to this strategy, however, is the requirement for the seemingly indispensable *N*-acyl group on the β -enamino esters; this group is needed for metal chelation, which improves reactivity and selectivity. The introduction of this moiety often produces enamine alkene isomers that can be difficult to separate, and, importantly, the individual isomers are typically hydrogenated with differing rates and selectivities. Moreover, these difficulties are magnified by the necessary removal of this group, a seemingly cumbersome artefact of an otherwise powerful strategy. Nonetheless, this advance has allowed a variety of β -amino acids to be prepared³¹.

An innovative solution to this problem was demonstrated by a group at Merck en route to synthesizing Januvia (sitagliptin phosphate; **8**, Fig. 2), which has recently been approved by the US Food and Drug Administration for the treatment of type 2 diabetes³⁵. The optimal target contains an unfunctionalized β -amino amide. A strategy was sought to install this moiety directly by asymmetric hydrogenation of unsubstituted β -enamino ester and amide derivatives³⁶ (for example, **4**, Fig. 2). A traditional hydrogenation route for the production of amino acids is a proven, cost-effective method for the synthesis of chiral building blocks. The industrial infrastructure is already in place to realize this goal; however, in this case, the reduction of unprotected β -enamino acids was not effective with existing chiral catalysts. A crucial component in addressing such limitations was Merck's high-throughput screening facility, which allowed rapid screening of catalyst structures and reaction conditions (an essential component for the success of any asymmetric catalytic process)³⁷. One potential complication for this hydrogenation strategy was avoided when it was observed that the preparation of the β -enamino ester and amide substrates (for example,

3→**4**, Fig. 2) proceeded with complete selectivity for the *Z*-isomer, presumably owing to hydrogen bonding in the products.

During the screening, a survey of transition metals and ligands revealed that rhodium complexes of the Josiphos (for example, **5**, Fig. 2) family of ligands efficiently catalyse the hydrogenation of a variety of substrates to give high yields with excellent enantioselectivities. The remarkable functional-group tolerance of this catalyst allowed the strategic implementation of this asymmetric transformation as the penultimate step of the synthesis, thereby maximizing the usefulness of the process and materials. Thus, phenylacetic acid derivative **1** was converted into β -ketoamide **3** in a one-pot procedure via acylation of Meldrum's acid, followed by treatment with triazole salt **2** (ref. 38) (Fig. 2). Exposure to ammonium acetate converted this into β -enamino amide **4** as a single enamine isomer. Hydrogenation of **4** in the presence of 0.30 mol% of rhodium(I) and ligand **5** provided β -amino amide **7** in >95% conversion and 95% enantiomeric excess. Subsequent recrystallization and salt formation with phosphoric acid gave Januvia (**8**, Fig. 2). Efforts to optimize efficiency and examine the mechanism of the asymmetric process revealed that reactivity and selectivity were dependent on the pH of the reaction solution³⁹. It was found that ~1 mol% of a mild acid (that is, ammonium chloride) was essential for the reaction to proceed reproducibly on a large scale. In addition, it was observed that hydrogenation of a related substrate under identical conditions with a deuterium gas atmosphere resulted in deuterium incorporation at the β -position only, suggesting that an imine is an intermediate (**6**, Fig. 2) and that an enamine–imine tautomerization process plays an important part in the mechanism³⁶. Interestingly, intermediates such as **6** have a striking similarity to asymmetric β -carbonyl hydrogenations pioneered by Ryoji Noyori⁴⁰.

This example demonstrates the development of asymmetric catalysis into a state-of-the-art science through maximizing the efficiency by minimizing unnecessary functionality, by using atom economy and by using extremely active catalysts. Moreover, the development of the catalyst system for the synthesis of Januvia exemplifies the continued need for subtly different catalysts to meet new synthetic demands. Building on the experience obtained during the development of a highly efficient enamide reduction towards α -amino acids, such large-scale industrial synthesis of important β -amino acids has been a relatively rapid process.

Fluivirucine A₁

Transition-metal-catalysed cross-coupling reactions (see page 314) have been used extensively for constructing C–C bonds and, consequently, have had a substantial effect on the field of complex molecule synthesis^{41,42}. The predominance of palladium and nickel catalysts in

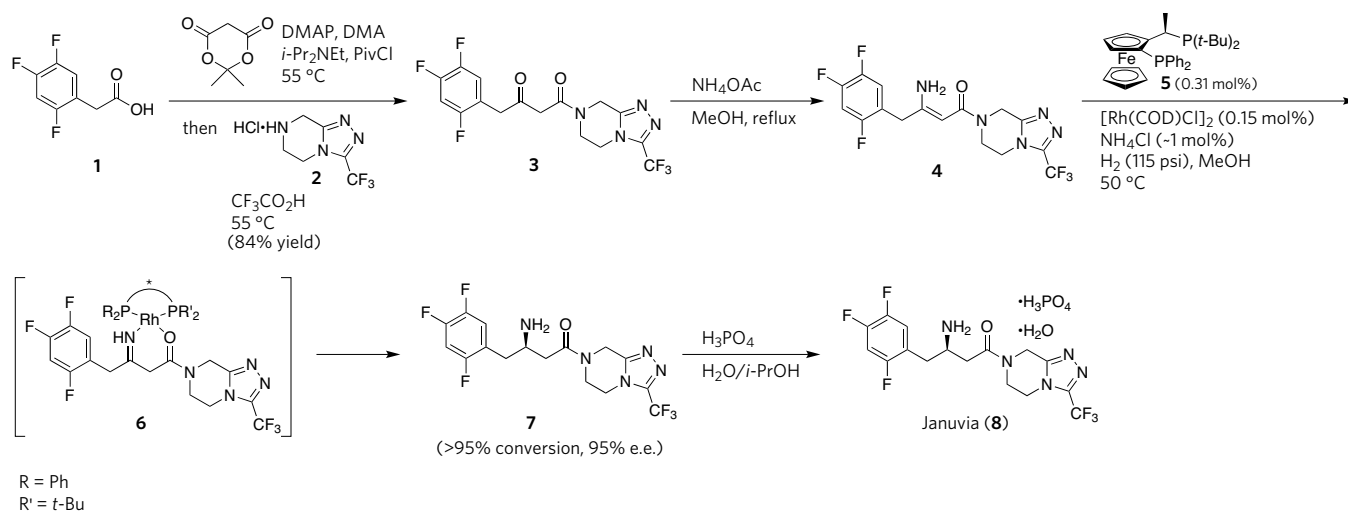


Figure 2 | Asymmetric hydrogenation of β -enamino amides towards Januvia. The rapid and selective synthesis of β -enamino amide (**3**) allowed a key rhodium•**5**-catalysed enantioselective hydrogenation under mildly acidic conditions, directly revealing the β -amino amide in Merck's synthesis

of Januvia (sitagliptin phosphate, **8**), an FDA-approved treatment for type 2 diabetes^{36,39}. COD, cyclooctadiene; DMA, dimethylacetamide; DMAP, 4-dimethylaminopyridine; e.e., enantiomeric excess; *i*-Pr, isopropyl; Piv, pivaloyl.

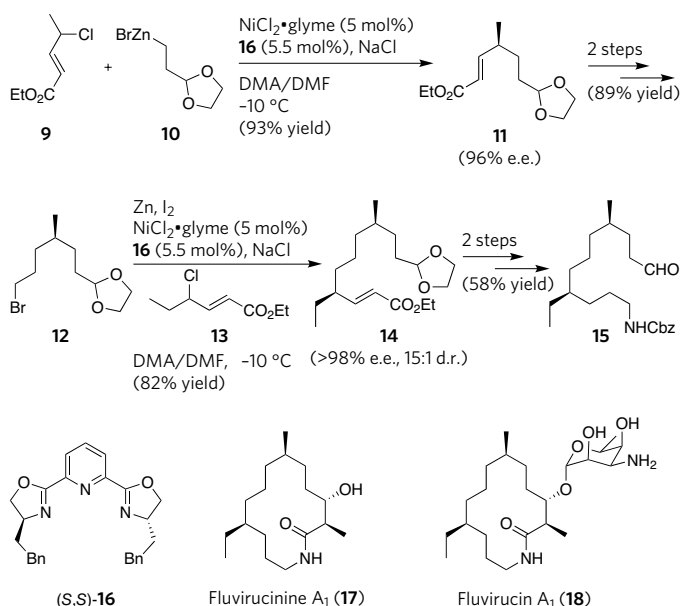


Figure 3 | Enantioselective $C(sp^3)$ - $C(sp^3)$ cross-coupling towards fluvirucine A_1 . Sequential asymmetric $C(sp^3)$ - $C(sp^3)$ Negishi cross-couplings of racemic allylic chlorides and alkylzinc reagents catalysed by nickel-(S,S)-16 enabled the rapid formal synthesis of fluvirucine A_1 (17), the aglycon of the antibiotic fluvirucine A_1 (18), with excellent enantioselectivity and diastereoselectivity, highlighting a creative solution to remote stereochemical control in unfunctionalized systems⁵¹. Cbz, benzyloxycarbonyl; DMF, dimethylformamide; d.r., diastereomeric ratio; glyme, 1,2-dimethoxyethane.

cross-coupling technologies and their extraordinary functional-group tolerance increases the efficiency of this process by allowing a large degree of functionalization before coupling. Moreover, the efficacy of this cross-coupling strategy for streamlining synthesis has allowed retrosynthetic analyses that had been thought impossible with standard, non-metal reactions. Until recently, however, most cross-coupling methods involved $C(sp^2)$ - $C(sp^2)$ or $C(sp^2)$ - $C(sp)$ centres, limiting the application potential. Two crucial issues associated with expanding the substrate scope to include $C(sp^3)$ - $C(sp^3)$ couplings are the relatively low reactivity of alkyl halides towards oxidative addition and the propensity of σ -alkyl organometallic complexes to undergo rapid β -hydrogen-elimination reactions^{43,44}. Practical solutions to this problem were first presented by Akira Suzuki and Paul Knochel, followed more recently by Greg Fu and others (see refs 44 and 45 for reviews). In general, the reaction scope now encompasses a variety of primary and secondary halides and pseudohalides as the electrophilic component, with organoboranes, boronic acids, alkylmagnesium halides and alkylzinc halides as the nucleophilic component⁴⁴. Although perhaps not developed in the context of a particular target molecule, progress in these cross-coupling methods has allowed retrosynthetic disconnections that were not practical previously. Asymmetric cross-coupling protocols could, in turn, allow the direct formation of remote stereocentres in relatively unfunctionalized molecules.

Early examples of catalytic asymmetric cross-coupling reactions involving $C(sp^3)$ - $C(sp^2)$ centres were explored by Makoto Kumada and co-workers in the late 1970s and produced moderate enantioselectivities^{46,47}. Despite these initial reports and the subsequent evolution of cross-coupling methods and asymmetric catalysis, a deficiency in the development of catalytic asymmetric methods for $C(sp^3)$ - $C(sp^3)$ couplings existed until Fu and co-workers^{48,49} reported an asymmetric Negishi coupling in 2005. Before this report, researchers in the Fu laboratory observed the proficiency of tridentate pybox ligands (for example, 16, Fig. 3) at enabling the room-temperature, nickel-catalysed Negishi coupling of symmetric secondary alkyl bromides and iodides⁵⁰. It was postulated that the tridentate nature of pybox ligands prevented

the undesired β -hydrogen-elimination pathway, which would require a vacant coordination site. Reaction optimization facilitated the development of several asymmetric variations that generate challenging stereocentres applicable to complex molecule synthesis, as demonstrated in Fu's formal total synthesis of fluvirucine A_1 (17, Fig. 3), the aglycon of the macrolactam antibiotic fluvirucine A_1 (ref. 51) (18). A key nickel(II)-catalysed asymmetric cross-coupling of racemic allylic chloride (9) and alkylzinc reagent (10) in the presence of (S,S)-16 generated γ -disubstituted enone (11) in an excellent yield and 96% enantiomeric excess. Elaboration over two steps to a bromide (12), followed by conversion to the alkylzinc form and a second nickel(II)-catalysed asymmetric Negishi cross-coupling with racemic allylic chloride 13, provided the ester 14 in a good yield and with >98% enantiomeric excess and a 15:1 diastereoisomer ratio. A subsequent two-step conversion to the aldehyde 15 intersected Young-Ger Suh's⁵² synthesis of fluvirucine A_1 (17). This method exemplifies the efficiency of the $C(sp^3)$ - $C(sp^3)$ cross-coupling and presents a creative solution to the particularly difficult challenge of remote stereochemical control.

At present, most examples of this technology require a stabilizing group adjacent to the site of the putative carbon-centred radical. Eliminating this condition would further improve the utility of this asymmetric cross-coupling method. In addition, stereogenic organometallic coupling partners (for example, secondary alkylzinc reagents) have not yet been reported in this asymmetric transformation. A potential goal for this synthetic method would be the combination of a racemic secondary alkyl halide and a racemic secondary alkylmetal reagent to form vicinal stereocentres along an alkyl chain with high levels of enantioselectivity and diastereoselectivity.

Minfiensine

The asymmetric generation of all-carbon quaternary stereocentres is a considerable challenge for synthetic chemists⁵³. As quaternary stereocentres are found in many natural product structures, convenient enantioselective methods for their formation would be useful. One such method is the Heck reaction⁵⁴, in which a palladium(0) catalyst promotes the vinylation of an aryl halide, vinyl halide or trifluoromethanesulphonate. The large body of literature on palladium catalysis and mechanisms⁴¹, as well as an ever-growing collection of chiral ligands for transition-metal catalysis, greatly increased the potential of using this method to carry out asymmetric catalysis. In addition, many synthetic endeavours using diastereoselective or non-stereoselective intramolecular Heck reactions have been reported (see ref. 55 for representative examples), increasing the significance of an enantioselective process. In 1989, the laboratories of Masakatsu Shibasaki⁵⁶ and Larry Overman⁵⁷ independently reported the first variants of an intramolecular catalytic, asymmetric Heck reaction. Initial levels of enantioselectivity were moderate; however, subsequent optimizations

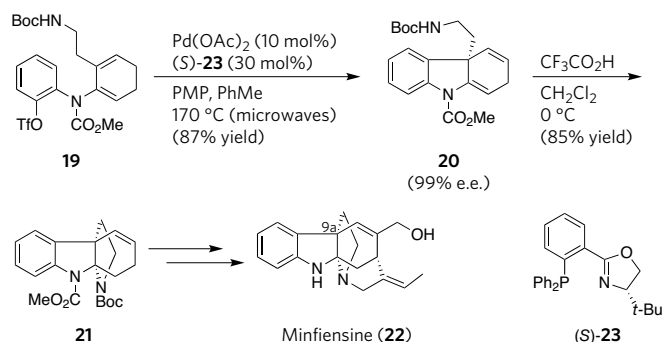


Figure 4 | Asymmetric Heck cyclization towards minfiensine. The palladium-23-catalysed enantioselective intramolecular Heck cyclization of 19 forged the C9a all-carbon quaternary stereocentre and, on acid-mediated cyclization, the polycyclic core of minfiensine (22). This method allowed the diastereoselective preparation of the remaining stereocentre and completion of the alkaloid⁶¹. Boc, *tert*-butoxycarbonyl; OTf, trifluoromethanesulphonate; PMP, 1,2,2,6,6-pentamethylpiperidine.

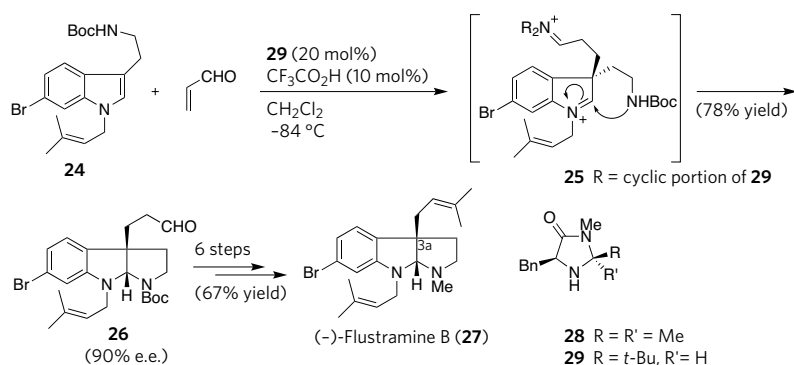


Figure 5 | Amine-catalysed indole alkylation towards flustramine B. The enantioselective C3a alkylation of the tryptamine derivative **24** catalysed by the amine **29**, on subsequent intramolecular cyclization, provided the pyrroloindoline core of (–)-flustramine B (**27**), a potassium-channel blocking agent. This selective cascade process prepared the all-carbon quaternary stereocentre, rapidly and efficiently enabling completion of flustramine B⁷⁴.

realized good-to-excellent selectivities in the generation of tertiary and all-carbon quaternary stereocentres⁵⁸.

Indole alkaloids encompass a large number of natural and pharmaceutical substances with a wide range of biological activities⁵⁹. The plant alkaloid minfiensine (**22**, Fig. 4) is a compelling example of the all-carbon quaternary stereocentre motif in biologically active natural products. Minfiensine and related alkaloids have been used in traditional medicines and have promising anticancer activity⁶⁰. The intriguing polycyclic structure and biological relevance of minfiensine prompted the Overman laboratory⁶¹ to explore a catalytic enantioselective Heck reaction to generate the sole quaternary stereocentre at C9a. It was discovered that the palladium-catalysed intramolecular Heck reaction of dienyl aryl trifluoromethanesulphonate **19** (Fig. 4) in the presence of the phosphinooxazoline ligand (*S*)-**23** (Fig. 4) under microwave conditions produced indoline **20** (Fig. 4) at a good yield and with 99% enantiomeric excess. Subsequent acid-promoted carbamate cyclization produced the tricyclic core of minfiensine (**21**, Fig. 4), which was then converted to the natural product. The efficiency and selectivity of the catalytic asymmetric Heck reaction facilitated completion of the target, where the remaining stereocentres are derived from this initial transformation.

Despite numerous examples of the asymmetric Heck reaction in total synthesis⁵⁸, there are several features that could be improved. Reactions typically require high temperatures and relatively high catalyst loadings, and the development of chiral ligands that greatly increase the reactivity of the transition metal while maintaining an adequate asymmetric environment would be greatly beneficial. As most enantioselective Heck reactions use an sp^2 -hybridized organohalide component, another frontier lies in the application of unactivated alkyl carbon electrophiles that have β -hydrogen in both intramolecular and intermolecular cases, an area currently in its infancy⁶².

Flustramine B

Numerous methods have been developed for the generation of substituted indoles^{63,64}; however, enantioselective indole functionalization has been far less explored. To address the deficiencies in the indole functionalization literature, Karl Anker Jørgensen⁶⁵ and David MacMillan⁶⁶ independently developed strategies for asymmetric Friedel–Crafts alkylation of conjugate acceptors with electron-rich heteroaromatics. MacMillan's method uses a secondary amine catalyst (**28**, Fig. 5) that facilitates the LUMO (lowest-unoccupied molecular orbital)-lowering activation of α,β -unsaturated aldehydes for a variety of transformations^{67,68} (see page 304). Although imidazolidinone **28** (Fig. 5) was a sufficient catalyst for the Friedel–Crafts alkylation of pyrroles, generating good yields and excellent enantioselectivities⁶⁶, application of less-activated indole substrates resulted in sluggish reactivity with considerably diminished selectivities⁶⁹. Kinetic investigations of iminium-catalysed reactions revealed that the overall reaction rate was influenced by the efficiency of formation for both the iminium ion and the C–C bond, prompting the development of a modified imidazolidinone catalyst (**29**, Fig. 5). This refinement minimized the steric bulk around one face of the catalyst, thereby exposing the lone pair of electrons on the secondary amine nitrogen. This structural

change translated into increased reactivity that enabled the asymmetric Friedel–Crafts alkylation of a variety of indoles with good-to-excellent yields and very high enantioselectivities⁶⁹.

Pyrroloindoline alkaloids are a family of polyindole alkaloids of diverse structural complexity and biological relevance⁷⁰. Diastereoselective syntheses of the core of these compounds have focused on the control of the C3a all-carbon quaternary stereocentre as a key design element^{71–73}. With a powerful and mild indole alkylation method in hand, MacMillan and co-workers devised a cascade strategy for the catalytic asymmetric preparation of the C3a stereocentre and the pyrroloindoline core of the potassium-channel blocker (–)-flustramine B (**27**, Fig. 5) in one step⁷⁴. In this key transformation, tryptamine derivative **24** (Fig. 5) and 2-propenal (acrolein), in the presence of catalyst **29**, underwent the asymmetric Friedel–Crafts alkylation to provide iminium intermediate **25** (Fig. 5). Subsequent carbamate cyclization and hydrolysis to regenerate the catalyst provided the core (**26**, Fig. 5) with a good yield and 90% enantiomeric excess. Importantly, this allowed completion of (–)-flustramine B (**27**) in just six steps and with good overall yield, highlighting the efficiency of this cascade approach. It is noteworthy that this strategy also has the potential to be applied to the synthesis of various polycyclic indolines such as the diazonamide family of cytotoxic alkaloids⁷⁴. It is also interesting to note that both the intramolecular Heck reaction (see the subsection ‘Minfiensine’) and the indole Friedel–Crafts alkylation can generate similar indoline structural motifs despite the markedly different bond-connecting strategies of these reactions. The success of

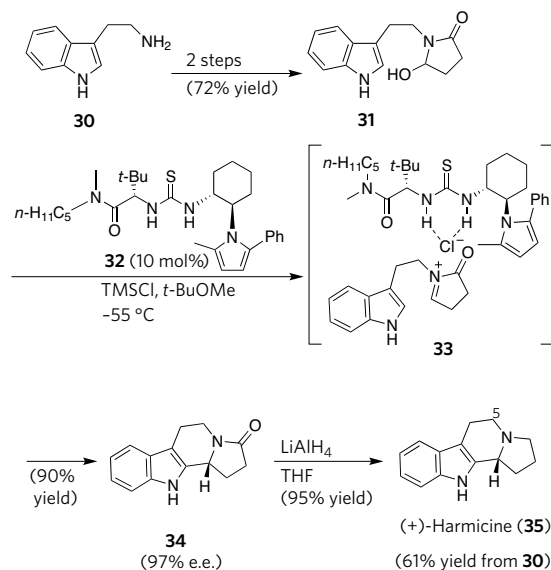


Figure 6 | Enantioselective Pictet–Spengler cyclization towards harmicine. Facile preparation of the hydroxylactam 31 aided asymmetric Pictet–Spengler cyclization catalysed by thiourea 32. This enantioselective *N*-acyl-iminium ion cyclization allowed rapid construction of the alkaloid (+)-harmicine (35) following the Pictet–Spengler cyclization⁸¹. THF, tetrahydrofuran.

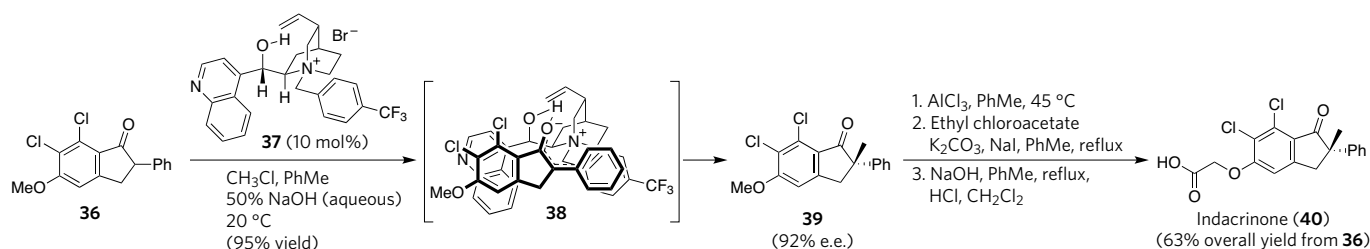


Figure 7 | Asymmetric phase-transfer alkylation towards indacrinone. Scientists at Merck developed an enantioselective alkaloid-salt-catalysed phase-transfer enolate alkylation to gain access to indanone **39**, en route to indacrinone (**40**) (ref. 84), a candidate diuretic drug.

these dissimilar strategies allows a great deal of flexibility in the planning of syntheses.

Iminium-activation methods with chiral amine catalysts have been successful for numerous transformations, but catalyst loading, turnover frequency and excesses of certain reagents limit the large-scale industrial application of these methods. In addition, in some cases, the organic catalyst may be more difficult to remove from the reaction products than a metal catalyst. However, the typically air- and moisture-stable reaction conditions, low cost of some catalysts and often metal-free conditions are attractive. The variety of asymmetric transformations (some proceeding through substantially different reaction pathways) that have been realized with chiral amine catalysts so far indicates a burgeoning field in which there are many useful enantioselective catalysts.

Harmicine

Since the intramolecular cyclization of an aromatic ring onto an iminium species was reported by Amé Pictet and Theodor Spengler⁷⁵ in 1911, this transformation has been of great use in the synthesis of many important alkaloid natural products⁷⁶. Indeed, the need for asymmetric variants of this reaction was recognized, and several diastereoselective protocols have been devised⁷⁶. A common approach to diastereoselective Pictet–Spengler cyclization has been to use tryptophan derivatives to control the stereochemistry of the cyclization. However, using this type of technique for the synthesis of a natural product such as harmicine (**35**, Fig. 6), which is active against the disease leishmaniasis, necessitates the removal of the stereocontrol element at C5, following the diastereoselective cyclization. Nonetheless, Steven Allin and co-workers⁷⁷ proved this to be a viable method in 2007. This particular structure, however, highlighted a challenge for enantioselective catalysis and an opportunity to improve synthetic efficiency.

When considering prospects for asymmetric induction, Eric Jacobsen and Mark Taylor considered activated *N*-acyl-iminium ions as a template and reasoned that a chiral thiourea derivative might be effective in promoting cyclization⁷⁸. In practice, these Brønsted acids⁷⁸, as well as other Brønsted acids investigated later by other groups^{79,80}, proved to be excellent catalysts for enantioselective indole annulations with *in situ*-generated *N*-acyl-iminium species (for example, **33**, Fig. 6). In later studies by Jacobsen and co-workers, it was found that hydroxylactams (for example, **31**, Fig. 6) are convenient precursors to *N*-acyl-iminium ions, which in turn enable access to various polycyclic structures^{81,82}. Given this effective protocol, an efficient catalytic asymmetric synthesis of harmicine (**35**) was realized in four steps from tryptamine (**30**, Fig. 6). Several mechanistic experiments have suggested that asymmetric induction is controlled by a complex of the Brønsted acid catalyst (**32**, Fig. 6) and a chloride counterion closely associated with the iminium ion (for example, **33**, Fig. 6) that effectively blocks approach to one face of the electrophile, providing annulated products (for example, **34**, Fig. 6) with excellent enantiomeric excesses. This insight into the remarkable mechanism of this transformation has led to a related C–C bond-forming process using oxocarbenium ions⁸³. Further exploitation of this unusual proposed catalyst–anion interaction could lead to a variety of other asymmetric addition reactions, such as intermolecular alkylation of *N*-acyl-iminium ions. In common with the history of the Diels–Alder reaction (see the section ‘Historical overview of enantioselective methods’), the exploration of the

Pictet–Spengler cyclization has provided a useful method to access many heterocyclic structures embedded in alkaloid natural products using a classical reaction with well-established synthetic applications.

Indacrinone

Enolate alkylations exemplify the fundamental usefulness of the carbonyl group for C–C bond formation. Strategies to induce asymmetry in these reactions have included chiral auxiliaries and chiral ligands, although few examples are catalytic. A particularly challenging class of product targets are all-carbon quaternary stereocentres adjacent to carbonyl groups. One example of an important target bearing this motif is the diuretic drug candidate indacrinone⁸⁴ (**40**, Fig. 7). Given the lack of efficient methods for synthesizing this structure, researchers at Merck envisaged an enantioselective phase-transfer alkylation method based on a quaternary ammonium salt derived from a naturally occurring cinchona alkaloid (for example, **37**, Fig. 7). In the event, readily prepared indanone **36** (Fig. 7) was methylated, producing **39** with 95% yield and 92% enantiomeric excess, and **39** was then converted to indacrinone (**40**, Fig. 7) in three additional steps. Although successful in achieving enantioselective enolate alkylation, the mechanism for this process seems to be complex⁸⁵; however, enantiofacial selectivity in the alkylation event may be rationalized through a hypothetical transition state (**38**, Fig. 7).

Three key interactions are thought to control selectivity: a hydrogen bond between the enolate oxygen and the catalyst hydroxyl group, and two π -system stacking interactions between the four aromatic rings. Perhaps as a consequence of the complex mechanism, the range of substrates for enolate alkylation is limited, and other solutions to this problem are still needed. However, these initial results have led to several related catalytic enantioselective reactions using cinchoninium salts or related organic ammonium complexes as catalysts⁸⁶. The discovery of these useful catalysts has provided not only an alternative to related transformations using metal catalysts but also a means of accessing chiral environments that are simply not possible with metal-based catalysts. Moreover, eliminating metal waste materials is attractive from an industrial and environmental standpoint. Ultimately, the studies directed towards an enantioselective synthesis of indacrinone demonstrate the versatility of privileged catalysts developed for the synthesis of target molecules for a range of other applications.

Cyanthiwigin F

A recent case of enantioselective enolate alkylation is the synthesis of cyanthiwigin F (**48**, Fig. 8), a cytotoxic natural product from a sea sponge. The cyanthiwigin family is composed of more than 30 diterpenoids, most of which bear two quaternary stereocentres, at C6 and C9, and a *syn* relationship of the methyl groups in the central ring. These central stereochemical elements are a complicating factor for a convergent strategy that might seek to couple the five- and seven-membered ring portions and subsequently form the six-membered ring. To avoid this difficulty, John Enquist and Brian Stoltz chose instead to address these two central stereocentres at an early stage and append the five- and seven-membered rings to the assembled cyclohexane. Accordingly, a synthetic strategy was devised that involved a one-pot, double-enantioselective enolate alkylation reaction to form both quaternary

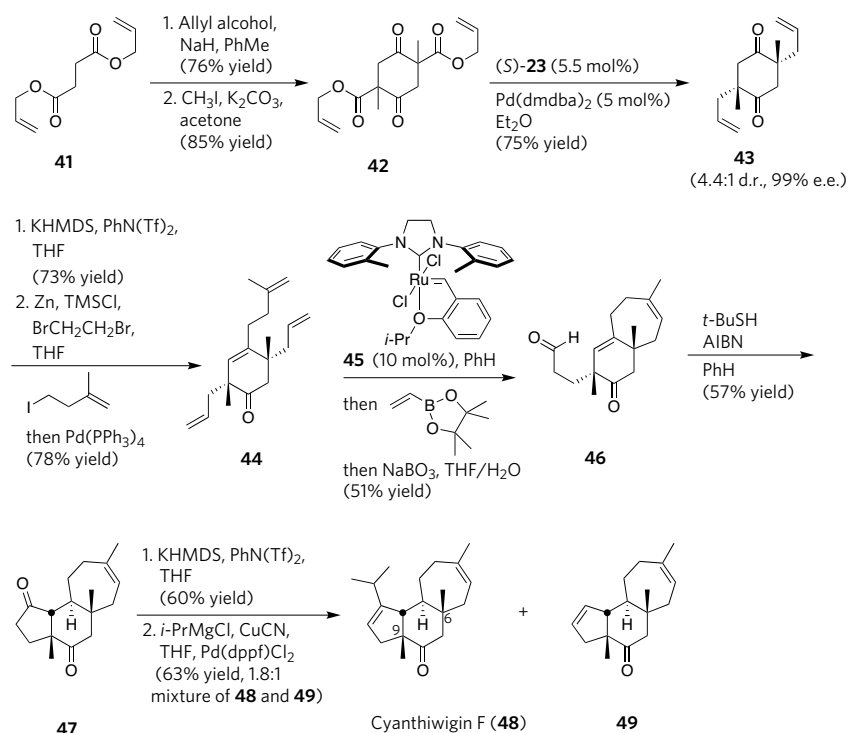


Figure 8 | Double enantioselective enolate alkylation and the synthesis of cyanthiwigin F. The preparation of a diastereomeric mixture of racemic and meso β-ketoester 42 allowed the strategic application of the palladium-(S)-23-catalysed double enantioselective enolate alkylation to generate both all-carbon quaternary stereocentres of 43 with excellent stereoselectivity. Subsequent Negishi coupling, ring-closing metathesis and radical-mediated aldehyde–olefin cyclization were used to construct the tricyclic core of the cytotoxic natural product cyanthiwigin F (48) rapidly, completing a synthesis involving several steps at which multiple C–C bonds are formed⁸⁷. AIBN, 2,2'-azobis(isobutyronitrile); dppf, 1,1'-bis(diphenyl-phosphino)ferrocene; dmdba, bis(3,5-dimethoxybenzylidene)acetone; KHMDS, potassium bis(trimethylsilyl)amide; PhN(Tf)₂, phenyl bis(trifluoromethane)sulphonimide.

stereocentres simultaneously. Although such enantioselective alkylations have proved difficult, recent studies have identified palladium catalysts that might provide a solution to this problem and enable the synthesis of a variety of targets containing quaternary carbon stereocentres, including the cyanthiwigins^{87,88}.

The implementation of this retrosynthetic strategy began with a Claisen–Dieckmann sequence that converted diallyl succinate (41, Fig. 8) to bis(β-ketoester) 42 (Fig. 8) as a 1:1 mixture of racemic and meso diastereoisomers. On exposure to the catalyst derived from Pd(dmdba)₂ and phosphinooxazoline ligand (S)-23 (refs 89, 90) (Fig. 4), each stereoisomer of 42 was transformed to bis(allylated) ketone 43 (Fig. 8) with 75% yield and 99% enantiomeric excess as a 4.4:1 mixture of diastereoisomers. With both quaternary centres in place, elaboration of this stereochemically rich core structure to the natural product was achieved in six further steps. Enol trifluoromethanesulphonate formation and Negishi coupling (43→44) preceded a tandem ring-closing metathesis–cross-metathesis sequence with Grubbs' ruthenium catalyst 45 (Fig. 8). Aldehyde–alkene radical cyclization generated the final ring of the cyanthiwigin core (47, Fig. 8), and enol trifluoromethanesulphonate formation and palladium-catalysed cross-coupling formed (–)-cyanthiwigin F (48, Fig. 8), together with reduction product 49. Choosing to confront the difficult stereochemical elements of the cyanthiwigin structure at an early stage led to a direct synthetic route proceeding in nine steps from diallyl succinate. This strategy was made possible by the intriguing reaction mechanism of the enantioselective decarboxylative allylation, in which all three stereoisomers of bis(β-ketoester) 42 are converted to a specific stereoisomer of product (43) with high selectivity, through a stereoablative process⁹¹. In addition, of the nine steps required for the synthesis, seven form C–C bonds, and four form multiple C–C bonds. Directly addressing the carbon framework of the target molecule and the stereochemical challenges embedded within ultimately led to an efficient synthetic sequence for this important molecule.

Recently, the proposed chiral palladium enolate was shown to be intercepted by allyl or proton electrophiles⁸⁸. Although the synthesis of cyanthiwigin F demonstrates the versatility of allyl moieties for further derivatization, the direct use of alternative electrophiles would provide a more general and direct method for transition-metal-mediated enolate functionalization.

Marcfortine B

Of the many fundamental approaches to the formation of five-membered rings from acyclic precursors, the [3 + 2] cycloaddition is among the most convergent strategies. A useful method of achieving such a cyclization is via a trimethylenemethane (TMM) intermediate⁹². This interesting non-Kekulé molecule was first prepared and studied through photolytic decomposition of a cyclic diazene precursor. However, the free diyl is prone to several undesired reaction pathways and does not lend itself to asymmetric catalysis. Despite this, intramolecular diyl-trapping reactions are valuable methods of cyclopentane formation⁹². Recognizing the synthetic utility of TMM, Barry Trost and co-workers developed an array of 2-(trimethylsilyl)-2-propenyl acetate reagents that generate a metal•TMM complex when exposed to a palladium catalyst^{93,94}. A recent application of this transformation in total synthesis is the approach to marcfortine B (55, Fig. 9a), a member of a family of anti-parasitic agents⁹⁵. The strategy used sought to forge the [2.2.2]bicyclic via an intramolecular radical cyclization and install the spiro all-carbon quaternary stereocentre by the cycloaddition of oxindole 50 (Fig. 9a) with TMM precursor 51 (Fig. 9a). In the event, an excellent yield was observed for the annulation reaction yielding spirooxindole 52 (Fig. 9a) as a 1:1 mixture of diastereoisomers. Over the course of nine additional steps, spirocycle 52 was transformed into amide 53 (Fig. 9a). Preparation of the xanthate derivative of alcohol 53 allowed radical cyclization, generating the challenging [2.2.2]bicycle 54 (Fig. 9a). Seven further steps produced (±)-marcfortine B (55, Fig. 9a).

Although this strategy demonstrated several intriguing ring-forming reactions, an asymmetric synthesis of 55 would require an enantioselective variant of the key TMM-[3 + 2] cycloaddition, a goal that has remained elusive⁹⁶. The first asymmetric palladium-catalysed [3 + 2] cycloaddition with various bis(phosphine) ligands was reported by Yoshihiko Ito and co-workers⁹⁷, but with only moderate enantiomeric excess (up to 78%) and diastereomeric ratio (up to 4:1 *trans:cis*). Thereafter, Trost and co-workers explored bulky monodentate phosphoramidite ligands (for example, (R,R,R)-59, Fig. 9b) for the transformation and observed very high enantioselectivity for the first time (refs 98–100) (Fig. 9b). Of particular interest is the enantioselective addition of substituted TMM reagents to functionalized oxindole derivatives⁹⁹. The use of oxindole 56 (Fig. 9b) and TMM-precursor 57 (Fig. 9b) in the palladium-catalysed cyclization with ligand (R,R,R)-59 yielded spirooxindole 58

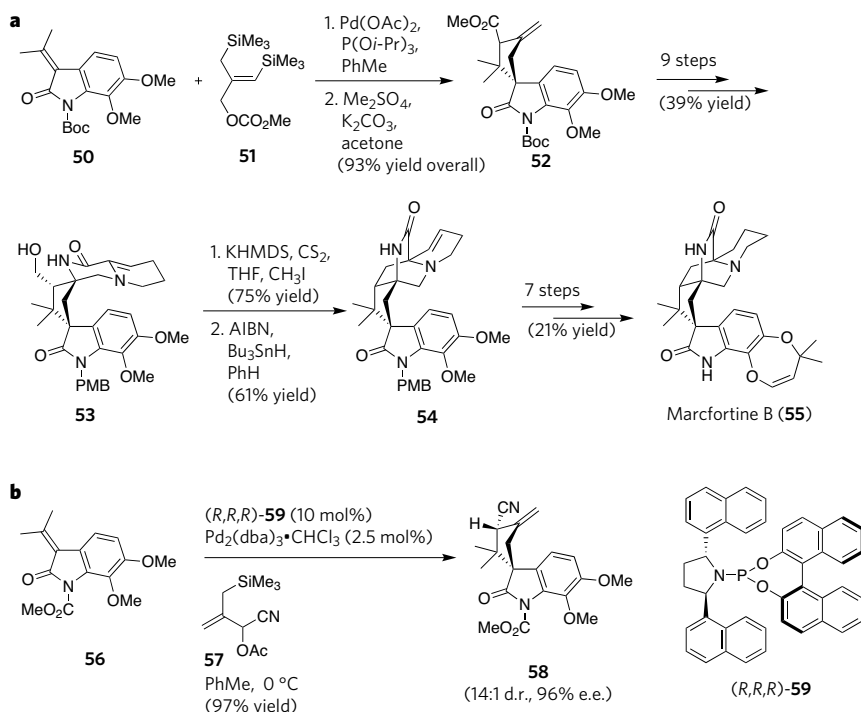


Figure 9 | Enantioselective TMM [3 + 2] cyclization towards marcfortine B. a, The palladium-catalysed TMM [3 + 2] cyclization of the oxindole 50 prepared the cyclopentane 52, giving an excellent yield. Elaboration to the tetracycle 53 facilitated an intramolecular radical cyclization that synthesized the core (54) of marcfortine B (55), which was completed in seven additional steps⁹⁵. **b**, The recent development of enantioselective palladium-catalysed TMM cyclizations with (*R,R,R*)-59 allows access to spiro-annulated products, such as 58, with high diastereoselectivity and enantioselectivity⁹⁹. The potential application of this method to synthesis of marcfortine B has not yet been realized. dba, bis(benzylidene)acetone; PMB, *p*-methoxybenzyl.

(Fig. 9b) with 14:1 diastereomeric ratio and 96% enantiomeric excess for the major diastereoisomer. Although a completed asymmetric synthesis of marcfortine B (55) from intermediate 58 has not been reported, many of the key functional groups are in place and the challenging spiro-quaternary stereocentre has been installed (compare 52 and 58). The development of this valuable asymmetric transformation highlights the ongoing efforts to devise new and useful techniques for the construction of important molecules.

Conclusions

The representative synthetic efforts presented here demonstrate the crucial interplay between target-directed synthesis and the development of novel reaction methods. Although many useful asymmetric technologies are currently available, the specific challenges posed by important natural products and pharmaceutical compounds highlight deficiencies in the current technology. Envisaging strategies to construct these relevant molecules through means beyond the current arsenal of enantioselective transformations will aid the evolution of both synthetic planning and reaction development. The symbiotic relationship between total synthesis and method development can continue to expand the understanding of synthetic strategy and catalysis on both fundamental and practical levels.

Despite the substantial advances that have been made so far, significant challenges remain for both multi-step synthesis and catalysis. In addition to improvements to efficiency and selectivity, better reactivity and handling stability are constantly required to implement and improve industrial processes for existing methods. Exceptionally reliable methods will aid in the discovery of new biologically active compounds by using high-throughput combinatorial screening techniques that are well established in the pharmaceutical industry, although these techniques are limited by the number of readily accessible chiral building blocks. Existing methods may be improved by identifying systems with better functional-group tolerances that might obviate the need for protecting and masking groups. Similarly, known privileged chiral frameworks may be modified to control chiral space more effectively for especially challenging transformations, a technique conspicuously successful for Trost's TMM cyclizations.

Overall, creative solutions are required to address specific organic transformations that remain significant impediments to efficient syntheses, namely forming multiple stereocentres and rings, forming

multiple C–C bonds, generating vicinal quaternary stereocentres and achieving C–H and C–C functionalization reactions. Cyclic structures often present particular challenges owing to the unique strain and steric elements imparted by their connectivity. As a result, many highly strained or complex polycyclic structures are daunting targets for synthesis. Finally, the discovery of new natural products will undoubtedly result in new challenges for synthetic chemistry and catalysis. ■

- Corey, E. J. & Cheng, X.-M. *The Logic of Chemical Synthesis* (Wiley, 1995).
- Jacobsen, E. N., Pfaltz, A. & Yamamoto, H. (eds) *Comprehensive Asymmetric Catalysis* Vol. I–III (Springer, 2000).
- Jacobsen, E. N., Pfaltz, A., Yamamoto, H. (eds) *Comprehensive Asymmetric Catalysis* Suppl. 1 & 2 (Springer, 2004).
- Hoveyda, A. H. in *Stimulating Concepts in Chemistry* (eds Vögtle, F., Stoddart, J. F. & Shibasaki, M.) 145–160 (Wiley-VCH, 2000).
- Trost, B. M. Asymmetric catalysis: an enabling science. *Proc. Natl Acad. Sci. USA* **101**, 5348–5355 (2004).
- Taylor, M. S. & Jacobsen, E. N. Asymmetric catalysis in complex target synthesis. *Proc. Natl Acad. Sci. USA* **101**, 5368–5373 (2004).
- Diels, O. & Alder, K. Synthesen in der Hydroaromatischen Reihe. *Justus Liebigs Ann. Chem.* **460**, 98–122 (1928).
- Stork, G., van Tamelen, E. E., Friedman, L. J. & Burgstahler, A. W. Cantharidin. A stereospecific total synthesis. *J. Am. Chem. Soc.* **73**, 4501 (1951).
- Woodward, R. B., Bader, F. E., Bickel, H., Frey, A. J. & Kierstead, R. W. The total synthesis of reserpine. *J. Am. Chem. Soc.* **78**, 2023–2025 (1956).
- Yates, P. & Eaton, P. Acceleration of the Diels–Alder reaction by aluminum chloride. *J. Am. Chem. Soc.* **82**, 4436–4437 (1960).
- Hashimoto, S.-i., Komeshima, N. & Koga, K. Asymmetric Diels–Alder reaction catalysed by chiral alkoxyaluminum dichloride. *J. Chem. Soc. Chem. Commun.* 437–438 (1979).
- Hayashi, Y. in *Cycloaddition Reactions in Organic Synthesis* (eds Kobayashi, S. & Jørgensen, K.) 5–55 (Wiley-VCH, 2002).
- Corey, E. J. Catalytic enantioselective Diels–Alder reactions: methods, mechanistic fundamentals, pathways, and applications. *Angew. Chem. Int. Edn Engl.* **41**, 1650–1667 (2002).
- Nicolaou, K. C., Snyder, S. A., Montagnon, T. & Vassilikogiannakis, G. The Diels–Alder reaction in total synthesis. *Angew. Chem. Int. Edn Engl.* **41**, 1668–1698 (2002).
- Corey, E. J. *et al.* Total synthesis of prostaglandins. Synthesis of the pure dL-E_1 , $\text{-F}_{1\alpha}$, $\text{-F}_{1\beta}$, -A_1 , and -B_1 hormones. *J. Am. Chem. Soc.* **90**, 3245–3247 (1968).
- Corey, E. J., Bakshi, R. K., Shibata, S., Chen, C.-P. & Singh, V. K. A stable and easily prepared catalyst for the enantioselective reduction of ketones. Applications to multistep syntheses. *J. Am. Chem. Soc.* **109**, 7925–7926 (1987).
- Corey, E. J. & Helal, C. J. Reduction of carbonyl compounds with chiral oxazaborolidine catalysts: a new paradigm for enantioselective catalysis and a powerful new synthetic method. *Angew. Chem. Int. Edn Engl.* **37**, 1986–2012 (1998).
- Yoon, T. P. & Jacobsen, E. N. Privileged chiral catalysts. *Science* **229**, 1691–1693 (2003).
- Nájera, C. & Sansano, J. M. Catalytic asymmetric synthesis of α -amino acids. *Chem. Rev.* **107**, 4584–4671 (2007).
- Senanayake, C. H. & Jacobsen, E. N. in *Process Chemistry in the Pharmaceutical Industry* (ed. Gadamasetti, K. G.) 347–368 (Marcel Dekker, 1999).

21. Akutagawa, S. & Tani, K. in *Catalytic Asymmetric Synthesis* (ed. Ojima, I.) 145–161 (Wiley-VCH, 2002).
22. Nugent, W. A., RajanBabu, T. V. & Burk, M. J. Beyond nature's chiral pool: enantioselective catalysis in industry. *Science* **259**, 479–483 (1993).
23. Farina, V., Reeves, J. T., Senanayake, C. H. & Song, J. J. Asymmetric synthesis of active pharmaceutical ingredients. *Chem. Rev.* **106**, 2734–2793 (2006).
24. Blaser, H.-U. & Schmidt, E. (eds) *Asymmetric Catalysis on Industrial Scale: Challenges, Approaches and Solutions* (Wiley-VCH, 2004).
25. Eder, U., Sauer, G. & Wiechert, R. New type of asymmetric cyclization to optically active steroid CD partial structures. *Angew. Chem. Int. Edn Engl.* **10**, 496–497 (1971).
26. Hajos, Z. G. & Parrish, D. R. Stereocontrolled total synthesis of 19-nor steroids. *J. Org. Chem.* **38**, 3244–3249 (1973).
27. Hajos, Z. G. & Parrish, D. R. Asymmetric synthesis of bicyclic intermediates of natural product chemistry. *J. Org. Chem.* **39**, 1615–1621 (1974).
28. Evans, D. A., Fitch, D. M., Smith, T. E. & Cee, V. J. Application of complex aldol reactions to the total synthesis of phorbazole B. *J. Am. Chem. Soc.* **122**, 10033–10046 (2000).
29. Palomo, C., Oiarbide, M. & Garcia, J. M. The aldol addition reaction: an old transformation at constant rebirth. *Chem. Eur. J.* **8**, 37–44 (2002).
30. Ma, J.-A. Recent developments in the catalytic asymmetric synthesis of α - and β -amino acids. *Angew. Chem. Int. Edn Engl.* **42**, 4290–4299 (2003).
31. Juaristi, E. & Soloshonok, V. A. (eds) *Enantioselective Synthesis of β -Amino Acids* (Wiley, 2005).
32. Vineyard, B. D., Knowles, W. S., Sabacky, M. J., Bachman, G. L. & Weinkauff, D. J. Asymmetric hydrogenation. Rhodium chiral bisphosphine catalyst. *J. Am. Chem. Soc.* **99**, 5946–5952 (1977).
33. Lubell, W. D., Kitamura, M. & Noyori, R. Enantioselective synthesis of β -amino acids based on BINAP–ruthenium(II) catalyzed hydrogenation. *Tetrahedr. Asymmetry* **2**, 543–554 (1991).
34. Drexler, H.-J. *et al.* Chiral β -amino acid derivatives via asymmetric hydrogenation. *Org. Process Res. Dev.* **7**, 355–361 (2003).
35. Drucker, D., Easley, C. & Kirkpatrick, P. Sitagliptin. *Nature Rev. Drug Discovery* **6**, 109–110 (2007).
36. Hsiao, Y. *et al.* Highly efficient synthesis of β -amino acid derivatives via asymmetric hydrogenation of unprotected enamines. *J. Am. Chem. Soc.* **126**, 9918–9919 (2004). **This paper describes the discovery of the rhodium-catalysed enantioselective hydrogenation of β -enamino amide and ester substrates.**
37. Shultz, C. S. & Krska, S. W. Unlocking the potential of asymmetric hydrogenation at Merck. *Acc. Chem. Res.* **40**, 1320–1326 (2007).
38. Xu, F. *et al.* Mechanistic evidence for an α -oxoketene pathway in the formation of β -ketoamides/esters via Meldrum's acid adducts. *J. Am. Chem. Soc.* **126**, 13002–13009 (2004).
39. Clausen, A. M. *et al.* Identification of ammonium chloride as an effective promoter of the asymmetric hydrogenation of β -enamine amide. *Org. Process Res. Dev.* **10**, 723–726 (2006). **This paper details efforts to scale up the enantioselective hydrogenation of a β -enamino amide towards the industrial-scale synthesis of the type 2 diabetes treatment sitagliptin (Januvia), as well as the impact of using ammonium chloride as an additive to maintain efficiency and selectivity in this process.**
40. Noyori, R., Kitamura, M. & Ohkuma, T. Toward efficient asymmetric hydrogenation: architectural and functional engineering of chiral molecular catalysts. *Proc. Natl Acad. Sci. USA* **101**, 5356–5362 (2004).
41. Negishi, E.-i. & de Meijere, A. (eds) *Handbook of Organopalladium Chemistry for Organic Synthesis* Vol. 1 (Wiley, 2002).
42. de Meijere, A. & Diederich, F. (eds) *Metal-Catalyzed Cross-Coupling Reactions* Vols 1 & 2 (Wiley-VCH, 2004).
43. Luh, T.-Y., Leung, M.-K. & Wong, K.-T. Transition metal-catalyzed activation of aliphatic C–X bonds in carbon–carbon bond formation. *Chem. Rev.* **100**, 3187–3204 (2000).
44. Frisch, A. C. & Beller, M. Catalysts for cross-coupling reactions with non-activated alkyl halides. *Angew. Chem. Int. Edn Engl.* **44**, 674–688 (2005).
45. Netherton, M. R. & Fu, G. C. in *Topics in Organometallic Chemistry: Palladium in Organic Synthesis* (Tsuiji, J. ed.) 85–108 (Springer, 2005).
46. Kiso, Y., Tamao, K., Miyake, N., Yamamoto, K. & Kumada, M. Asymmetric cross-coupling reactions of *sec*-alkyl Grignard reagents with organic halides in the presence of a chiral phosphine–nickel complex as a catalyst. *Tetrahedr. Lett.* **15**, 3–6 (1974).
47. Hayashi, T. *et al.* Asymmetric synthesis catalyzed by chiral ferrocenylphosphine–transition metal complexes. 2. Nickel- and palladium-catalyzed asymmetric Grignard cross-coupling. *J. Am. Chem. Soc.* **104**, 180–186 (1982).
48. Fischer, C. & Fu, G. C. Asymmetric nickel-catalyzed Negishi cross-couplings of secondary α -bromo amides with organozinc reagents. *J. Am. Chem. Soc.* **127**, 4594–4595 (2005).
49. Arp, F. O. & Fu, G. C. Catalytic enantioselective Negishi reactions of racemic secondary benzylic halides. *J. Am. Chem. Soc.* **127**, 10482–10483 (2005).
50. Zhou, J. & Fu, G. C. Cross-couplings of unactivated secondary alkyl halides: room-temperature nickel-catalyzed Negishi reactions of alkyl bromides and iodides. *J. Am. Chem. Soc.* **125**, 14726–14727 (2003).
51. Son, S. & Fu, G. C. Nickel-catalyzed asymmetric Negishi cross-couplings of secondary allylic chlorides with alkylzincs. *J. Am. Chem. Soc.* **130**, 2756–2757 (2008). **This paper describes the use of two sequential nickel-catalysed asymmetric Negishi cross-coupling reactions in the formal synthesis of fluvirucinine A₁.**
52. Suh, Y.-G. *et al.* Asymmetric total synthesis of fluvirucinine A₁. *Angew. Chem. Int. Edn Engl.* **38**, 3545–3547 (1999).
53. Douglas, C. J. & Overman, L. E. Catalytic asymmetric synthesis of all-carbon quaternary stereocenters. *Proc. Natl Acad. Sci. USA* **101**, 5363–5367 (2004).
54. Heck, R. F. Palladium-catalyzed reactions of organic halides with olefins. *Acc. Chem. Res.* **12**, 146–151 (1979).
55. Overman, L. E. Application of intramolecular Heck reactions for forming congested quaternary carbon centers in complex molecule total synthesis. *Pure Appl. Chem.* **66**, 1423–1430 (1994).
56. Sato, Y., Sodeoka, M. & Shibasaki, M. Catalytic asymmetric carbon–carbon bond formation: asymmetric synthesis of *cis*-decalin derivatives by palladium-catalyzed cyclization of prochiral alkenyl iodides. *J. Org. Chem.* **54**, 4738–4739 (1989).
57. Carpenter, N. E., Kucera, D. J. & Overman, L. E. Palladium-catalyzed polyene cyclizations of trienyl triflates. *J. Org. Chem.* **54**, 5846–5848 (1989).
58. Dounay, A. B. & Overman, L. E. The asymmetric intramolecular Heck reaction in natural product total synthesis. *Chem. Rev.* **103**, 2945–2963 (2003).
59. Saxton, J. E. (ed.) *Chemistry of Heterocyclic Compounds: Indoles, Part Four, The Monoterpenoid Indole Alkaloids* (Wiley, 1983).
60. Dounay, A. B., Overman, L. E. & Wroblewski, A. D. Sequential catalytic asymmetric Heck–iminium ion cyclization: enantioselective total synthesis of the *Strychnos* alkaloid minfiensine. *J. Am. Chem. Soc.* **127**, 10186–10187 (2005).
61. Dounay, A. B., Humphreys, P. G., Overman, L. E. & Wroblewski, A. D. Total synthesis of the *Strychnos* alkaloid (+)-minfiensine: tandem enantioselective intramolecular Heck–iminium ion cyclization. *J. Am. Chem. Soc.* **130**, 5368–5377 (2008). **This paper provides a full account of the use of the palladium-catalysed enantioselective Heck reaction for total synthesis of the alkaloid minfiensine.**
62. Kirmansjah, L. & Fu, G. C. Intramolecular Heck reactions of unactivated alkyl halides. *J. Am. Chem. Soc.* **129**, 11340–11341 (2008).
63. Sundberg, R. J. *Indoles* (Academic, 1996).
64. Humphrey, G. R. & Kuethe, J. T. Practical methodologies for the synthesis of indoles. *Chem. Rev.* **106**, 2875–2911 (2006).
65. Jensen, K. B., Thorhauge, J., Hazell, R. G. & Jørgensen, K. A. Catalytic asymmetric Friedel–Crafts alkylation of β , γ -unsaturated α -ketoesters: enantioselective addition of aromatic C–H bonds to alkenes. *Angew. Chem. Int. Edn Engl.* **40**, 160–163 (2001).
66. Paras, N. A. & MacMillan, D. W. C. New strategies in organic catalysis: the first enantioselective organocatalytic Friedel–Crafts alkylation. *J. Am. Chem. Soc.* **123**, 4370–4371 (2001).
67. Lelais, G. & MacMillan, D. W. C. Modern strategies in organic catalysis: the advent and development of iminium activation. *Aldrichimica Acta* **39**, 79–87 (2006).
68. Erkkilä, A., Majander, I. & Pihko, P. M. Iminium catalysis. *Chem. Rev.* **107**, 5416–5470 (2007).
69. Austin, J. F. & MacMillan, D. W. C. Enantioselective organocatalytic indole alkylations. Design of a new and highly effective chiral amine for iminium catalysis. *J. Am. Chem. Soc.* **124**, 1172–1173 (2002).
70. Hino, T. & Nakagawa, M. in *The Alkaloids* Vol. 34 (ed. Brossi, A.) 1–75 (Academic, 1988).
71. Overman, L. E., Paone, D. V. & Stearns, B. A. Direct stereo- and enantiocontrolled synthesis of vicinal stereogenic quaternary carbon centers. Total syntheses of *meso*- and (–)-chimonanthine and (+)-calycanthine. *J. Am. Chem. Soc.* **121**, 7702–7703 (1999).
72. Overman, L. E., Larrow, J. F., Stearns, B. A. & Vance, J. M. Enantioselective construction of vicinal stereogenic quaternary stereocenters by dialkylation: practical total syntheses of (+)- and *meso*-chimonanthine. *Angew. Chem. Int. Edn Engl.* **39**, 213–215 (2000).
73. Depew, K. M. *et al.* Total synthesis of 5-N-acetylardeemin and amauroamine: Practical routes to potential MDR reversal agents. *J. Am. Chem. Soc.* **121**, 11953–11963 (1999).
74. Austin, J. F., Kim, S.-G., Sinz, C. J., Xiao, W.-I. & MacMillan, D. W. C. Enantioselective organocatalytic construction of pyrrolindolines by a cascade addition–cyclization strategy: synthesis of (–)-flustramine B. *Proc. Natl Acad. Sci. USA* **101**, 5482–5487 (2004). **This paper describes the use of imidazolidinone catalysts for enantioselective Friedel–Crafts reactions with indole nucleophiles for the synthesis of the potassium-channel blocker flustramine B.**
75. Pictet, A. & Spengler, T. Über die Bildung von Isochinolin-derivaten durch Einwirkung von Methylal auf Phenyl-äthylamin, Phenyl-alanin und Tyrosin. *Ber. Dtsch. Chem. Ges.* **44**, 2030–2036 (1911).
76. Cox, E. D. & Cook, J. M. The Pictet–Spengler condensation: a new direction for an old reaction. *Chem. Rev.* **95**, 1797–1842 (1995).
77. Allin, S. M., Gaskell, S. N., Elsegood, M. R. J. & Martin, W. P. A new asymmetric synthesis of the natural enantiomer of the indolizidine [8,7-b]indole alkaloid (+)-harmicine. *Tetrahedr. Lett.* **48**, 5669–5671 (2007).
78. Taylor, M. S. & Jacobsen, E. N. Highly enantioselective catalytic acyl–Pictet–Spengler reactions. *J. Am. Chem. Soc.* **126**, 10558–10559 (2004).
79. Seayad, J., Seayad, A. M. & List, B. Catalytic asymmetric Pictet–Spengler reaction. *J. Am. Chem. Soc.* **128**, 1086–1087 (2006).
80. Wanner, M. J., van der Haas, R. N. S., de Cuba, K. R., van Maarseveen, J. H. & Hiemstra, H. Catalytic asymmetric Pictet–Spengler reactions via sulfonyliminium ions. *Angew. Chem. Int. Edn Engl.* **46**, 7485–7487 (2007).
81. Raheem, I. T., Thiara, P. S., Peterson, E. A. & Jacobsen, E. N. Enantioselective Pictet–Spengler-type cyclizations of hydroxylactams: H-bond donor catalysis by anion binding. *J. Am. Chem. Soc.* **129**, 13404–13405 (2007). **This paper describes development of the thiourea-catalysed enantioselective Pictet–Spengler-type cyclization reaction used in the synthesis of the compound harmicine, which is active against leishmaniasis.**
82. Raheem, I. T., Thiara, P. S. & Jacobsen, E. N. Regio- and enantioselective catalytic cyclization of pyrroles onto N-acyliminium ions. *Org. Lett.* **10**, 1577–1580 (2008).
83. Reisman, S. E., Doyle, A. G. & Jacobsen, E. N. Enantioselective thiourea-catalyzed additions to oxocarbenium ions. *J. Am. Chem. Soc.* **130**, 7198–7199 (2008).
84. Dolling, U.-H., Davis, P. & Grabowski, E. J. J. Efficient catalytic asymmetric alkylations. 1. Enantioselective synthesis of (+)-indacrinone via chiral phase-transfer catalysis. *J. Am. Chem. Soc.* **106**, 446–447 (1984). **This paper describes the use of cinchoninium catalysts for enantioselective enolate alkylation towards synthesizing the diuretic indacrinone.**
85. Hughes, D. L., Dolling, U.-H., Ryan, K. M., Schoenewaldt, E. F. & Grabowski, E. J. J. Efficient catalytic asymmetric alkylations. 3. A kinetic and mechanistic study of the enantioselective phase-transfer methylation of 6,7-dichloro-5-methoxy-2-phenyl-1-indanone. *J. Org. Chem.* **52**, 4745–4752 (1987).
86. Maruoka, K. (ed.) *Asymmetric Phase Transfer Catalysis* (Wiley-VCH, 2008).
87. Enquist, J. A. Jr & Stoltz, B. M. The total synthesis of (–)-cynanthiwigin F by means of double catalytic enantioselective alkylation. *Nature* **453**, 1228–1231 (2008).

- This paper details the use of palladium-catalysed enantioselective enolate alkylation for the total synthesis of the marine natural product cyanthiwigin F.
88. Mohr, J. T. & Stoltz, B. M. Enantioselective Tsuji allylations. *Chem. Asian J.* **2**, 1476–1491 (2007).
89. Behenna, D. C. & Stoltz, B. M. The enantioselective Tsuji allylation. *J. Am. Chem. Soc.* **126**, 15044–15045 (2004).
90. Mohr, J. T., Behenna, D. C., Harned, A. M. & Stoltz, B. M. Deracemization of quaternary stereocenters by Pd-catalyzed enantioconvergent decarboxylative allylation of racemic β -ketoesters. *Angew. Chem. Int. Edn Engl.* **44**, 6924–6927 (2005).
91. Mohr, J. T., Ebner, D. C. & Stoltz, B. M. Catalytic enantioselective stereobalancing reactions: an unexploited approach to enantioselective catalysis. *Org. Biomol. Chem.* **5**, 3571–3576 (2007).
92. Yamago, S. & Nakamura, E. [3 + 2] Cycloaddition of trimethylenemethane and its synthetic equivalents. *Org. React.* **61**, 1–217 (2002).
93. Trost, B. M. & Chan, D. M. T. New conjunctive reagents. 2-Acetoxyethyl-3-allyltrimethylsilane for methylenecyclopentane annulations catalyzed by palladium(0). *J. Am. Chem. Soc.* **101**, 6429–6432 (1979).
94. Trost, B. M. & Chan, D. M. T. 2-Acetoxyethyl-3-allyltrimethylsilane and palladium(0): a source of trimethylenemethane–palladium complex? *J. Am. Chem. Soc.* **101**, 6432–6433 (1979).
95. Trost, B. M., Cramer, N. & Bernsmann, H. Concise total synthesis of (\pm)-marcfortine B. *J. Am. Chem. Soc.* **129**, 3086–3087 (2007).
- This paper describes the synthetic route to (\pm)-marcfortine B.
96. Le Marquand, P. & Tam, W. Enantioselective palladium-catalyzed trimethylenemethane [3 + 2] cycloadditions. *Angew. Chem. Int. Edn Engl.* **47**, 2926–2928 (2008).
97. Yamamoto, A., Ito, Y. & Hayashi, T. Asymmetric [3 + 2] cycloaddition of 2-(sulfonylmethyl)-2-propenyl carbonate catalyzed by chiral ferrocenylphosphine–palladium complexes. *Tetrahedr. Lett.* **30**, 375–378 (1989).
98. Trost, B. M., Stambuli, J. P., Silverman, S. M. & Schwörer, W. Palladium-catalyzed asymmetric [3 + 2] trimethylenemethane cycloaddition reactions. *J. Am. Chem. Soc.* **128**, 13328–13329 (2006).
99. Trost, B. M., Cramer, N. & Silverman, S. M. Enantioselective construction of spirocyclic oxindolic cyclopentanes by palladium-catalyzed trimethylenemethane [3 + 2]-cycloaddition. *J. Am. Chem. Soc.* **129**, 12396–12397 (2007).
- This paper describes palladium-catalysed enantioselective TMM [3 + 2] cycloaddition with oxindoles, which might provide an enantioselective route to marcfortine B.
100. Trost, B. M., Silverman, S. M. & Stambuli, J. P. Palladium-catalyzed asymmetric [3 + 2] cycloaddition of trimethylenemethane with imines. *J. Am. Chem. Soc.* **129**, 12398–12399 (2007).

Acknowledgements We thank the National Institute of General Medical Sciences (grant number R01GM080269-01), Eli Lilly (which provided predoctoral fellowships to J.T.M. and M.R.K.), Amgen, Abbott Laboratories, Boehringer Ingelheim, Merck, Bristol-Myers Squibb and the California Institute of Technology for financial support.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to B.M.S. (stoltz@caltech.edu).

Biologically inspired oxidation catalysis

Lawrence Que Jr¹ & William B. Tolman¹

The development of processes for selective hydrocarbon oxidation is a goal that has long been pursued. An additional challenge is to make such processes environmentally friendly, for example by using non-toxic reagents and energy-efficient catalytic methods. Excellent examples are naturally occurring iron- or copper-containing metalloenzymes, and extensive studies have revealed the key chemical principles that underlie their efficacy as catalysts for aerobic oxidations. Important inroads have been made in applying this knowledge to the development of synthetic catalysts that model enzyme function. Such biologically inspired hydrocarbon oxidation catalysts hold great promise for wide-ranging synthetic applications.

The selective oxidation of organic molecules is fundamentally important to life and immensely useful in industry^{1,2}. Among the myriad ways by which such reactions may be performed, those catalysed by naturally occurring metalloenzymes using molecular oxygen (O₂) are notable for many reasons. Metalloenzyme-catalysed oxidations often exhibit exquisite substrate specificity as well as regioselectivity and/or stereoselectivity, and operate under mild conditions through inherently 'green' processes. Moreover, metalloenzymes are sometimes able to alter the function of recalcitrant substrates in ways that synthetic chemists find difficult to replicate (for example, changing methane to methanol)^{1,3}. For these reasons, combined with the knowledge that metal centres in proteins often adopt novel structures and exhibit unusual properties that are intrinsically worth examining, bioinorganic chemists have studied O₂-activating metalloenzyme structure–function relationships extensively.

Significant advances in our understanding of how these enzymes function have accrued, in part from the high-resolution structures of resting states and reactive metalloenzyme intermediates through X-ray crystallography and spectroscopic characterization, and from detailed mechanistic information from intertwined kinetics, synthetic modelling and theoretical investigations. Such studies have produced candidates for the active oxidants in several classes of important copper and iron enzymes that are the subject of this review. These findings have fuelled the debate about the mechanisms by which these enzymes operate; one common issue being whether O–O bond cleavage occurs before, during or after attack on the organic substrate.

Although valuable in its own right, an added benefit of knowing the metalloenzyme structure and function is its potential application in the design of synthetic catalysts. Such 'bioinspired' or 'biomimetic' catalysts^{4,5} may have an advantage over metalloenzyme systems, insofar as they might expand the scope of possible substrates, increase the scale of production and tune selectivity and/or specificity (for example, reversal of asymmetric induction in a stereoselective process). They would also be useful for environmentally friendly catalytic chemistry⁶, where it is important to avoid the use of toxic or expensive metal reagents and oxidants, energy-consuming processing steps and undesirable reaction media. Furthermore, mechanistic studies of biomimetic catalysts can provide important insights into biological pathways, thus completing a feedback loop relating studies of metalloenzymes to their synthetic models. Indeed, recent advances in the design of biologically inspired oxidation catalysts containing inexpensive and readily available iron and copper centres have led to a new understanding of the fundamental

reaction steps and reactive intermediates relevant to metalloenzymes that incorporate these metals in their active sites, as well as to practical applications. In this review, we survey several examples of these advances, with particular emphasis in each case on the interplay of catalyst design and knowledge of metalloenzyme structure and function.

Biochemical inspirations

Iron and copper ions are the metal ions of choice for many biological oxidations because of their abundance in the geosphere, inherent electronic properties and accessible redox potentials. Pertinent examples are enzymes that contain haem iron, non-haem iron and copper active sites.

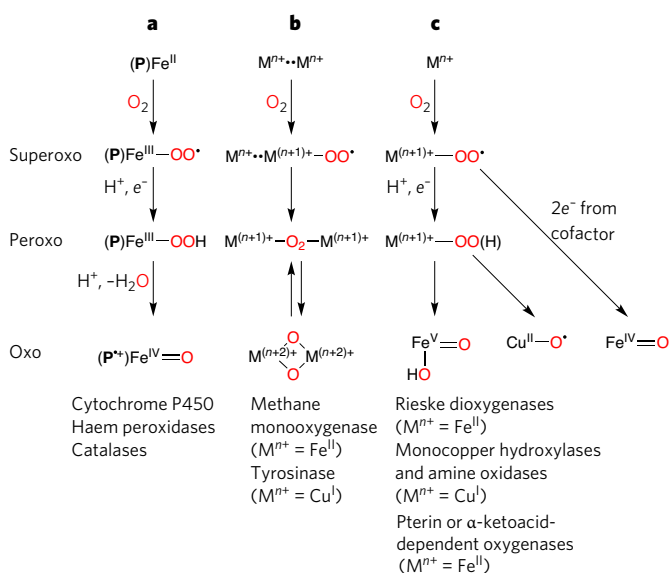


Figure 1 | Metallo-oxygenase mechanisms. To emphasize parallels with the haem paradigm (**a**), mechanisms proposed for O₂ activation by various metallo-oxygenases — di-iron and dicopper (**b**), and mononuclear non-haem iron and copper (**c**) — are shown. All involve the formation of an initial O₂ adduct (superoxo), conversion to a metal–peroxide (peroxo), and subsequent O–O bond cleavage to yield a high-valent oxidant (oxo). Oxygen atoms involved in these processes are shown in red. e[−], electron; M, metal; P, porphyrin.

¹Department of Chemistry and Center for Metals in Biocatalysis, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455, USA.

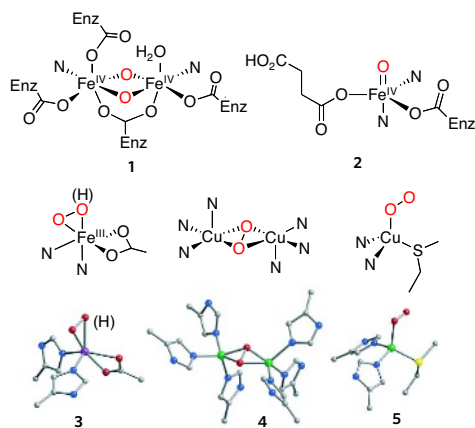


Figure 2 | Metalloenzyme intermediates. Diverse O_2 -derived metalloenzyme intermediates have been proposed on the basis of spectroscopic and theoretical investigations (1 and 2) and have been structurally characterized by X-ray crystallography (3–5). Species 1 is postulated to hydroxylate methane in soluble methane monooxygenase, whereas 2 is proposed to be the active oxidant in the non-haem iron α -ketoacid-dependent enzymes. Species 3 is an oxygenated intermediate from naphthalene dioxygenase; 4 is the $(\mu-\eta^2:\eta^2)$ -peroxodicopper(II) core of oxy-tyrosinase (akin to that identified in haemocyanin and catechol oxidase); and 5 is an O_2 adduct identified in peptidylglycine α -hydroxylating monooxygenase. For clarity, histidine imidazolyl donor ligands are indicated by N. Key oxygen atoms involved in these processes are shown in red. In the molecular models, carbon is shown in grey, oxygen in red, nitrogen in blue, copper in green, sulphur in yellow and iron in purple. Enz, enzyme.

The haem paradigm

The most extensively studied oxygen-activating enzymes are the cytochromes P450 (ref. 7). They carry out the hydroxylation of aliphatic C–H bonds and the epoxidation of C=C double bonds (C=C bonds) with high regioselectivity and stereoselectivity. Cytochromes P450 have an active site that consists of an iron porphyrin cofactor attached to the protein backbone through coordination of a cysteine at one of the axial positions on the metal, leaving the other axial position available for O_2 binding and activation. Related haem peroxidases and catalases activate peroxides with similar active sites where the axial cysteine is replaced by histidine or tyrosine⁸.

The activation of O_2 at a metal centre generally entails its two-electron reduction to the peroxo state and subsequent O–O bond cleavage.

The generally accepted oxygen activation mechanism associated with cytochrome P450 is referred to as the haem paradigm⁹ (Fig. 1a). In the initial step, O_2 coordinates to the reduced iron centre. It becomes progressively reduced to superoxo and peroxo forms and undergoes O–O bond cleavage to generate a formally oxoiron(V) oxidant that carries out the two-electron oxidation of the substrate. For haem enzymes, the two oxidizing equivalents required for oxidation are not stored at the iron centre but are instead delocalized on the iron porphyrin unit, so the formally oxoiron(V) oxidant in haem enzymes is generally described as an oxoiron(IV)-(oxidized porphyrin radical) species. Such delocalization allows access to a potent oxidant that is competent to attack a variety of substrates, a common strategy in biocatalysis that informs biomimetic design efforts.

Non-haem iron enzymes

As outlined in Fig. 1, the proposed mechanisms for non-haem iron enzymes in general follow the haem paradigm, and evidence for iron(III)–peroxo and high-valent iron–oxo intermediates has been obtained for some of these enzymes. For example, the di-iron enzyme methane monooxygenase (MMO)¹⁰ that catalyses the conversion of methane to methanol activates O_2 via di-iron(III)–peroxo and di-iron(IV) intermediates (Fig. 1b). Although the former has been implicated as the oxidant in the epoxidation of electron-rich alkenes such as ethyl vinyl ether¹¹, the latter has been demonstrated to be kinetically competent to hydroxylate methane¹². This di-iron(IV) species has been shown by extended X-ray absorption fine structure (EXAFS) analysis to have a $di(\mu-oxo)di-iron(IV)$ core¹³ (1, Fig. 2).

The Rieske dioxygenases, by contrast, activate O_2 at a mononuclear iron centre bound to a 2-His-1-carboxylate facial triad motif¹⁴ (Fig. 1c). These enzymes catalyse the *cis*-dihydroxylation of arene double bonds to initiate the biodegradation of aromatics in the soil. For these enzymes, the only redox centre that can store the two required oxidizing equivalents is the iron atom, so an oxoiron(V) oxidant has been invoked for this reaction. Although direct evidence for a non-haem oxoiron(V) species in the enzyme cycle has not yet been obtained, its iron(III)–peroxo precursor (3, Fig. 2) has been characterized by X-ray crystallography¹⁵. It is proposed that the side-on peroxo moiety either attacks the arene double bond directly or isomerizes first to an oxo(hydroxo)iron(V) species. Several examples of oxoiron(IV) intermediates (2, Fig. 2) have been trapped and characterized¹⁶ in studies of non-haem iron enzymes that use organic cofactors such as α -ketoglutarate or tetrahydrobiopterin, lending credence to the general mechanistic scheme presented in Fig. 1. Importantly, each of the side-on peroxo, oxoiron(V) or oxoiron(IV) species postulated for the non-haem iron enzymes is a viable target for the development of biologically inspired catalysts.

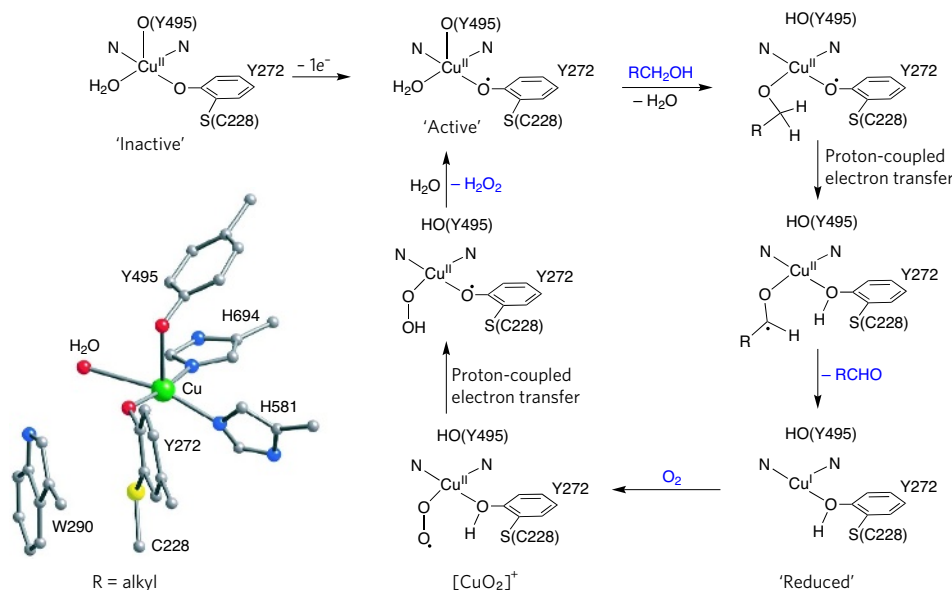


Figure 3 | Galactose oxidase mechanism.

The X-ray structure for the 'inactive' form of galactose oxidase (GAO) (Protein Data Bank identity 1GOG)²⁰ illustrates the disposition of ligands in the active site of the enzyme. Important components of the proposed mechanism for this enzyme include the involvement of the Cu(II)-tyrosyl radical unit ('active') in alcohol oxidation through intramolecular proton-coupled electron transfer. This unit is regenerated from a Cu(I) intermediate that reacts with O_2 to yield a $[CuO_2]^+$ species, which evolves H_2O_2 via another proton-coupled electron-transfer process. For clarity, histidine imidazolyl donor ligands are indicated by N. Organic substrates and products are shown in blue. In the molecular models, carbon is shown in grey, oxygen in red, nitrogen in blue, copper in green and sulphur in yellow.

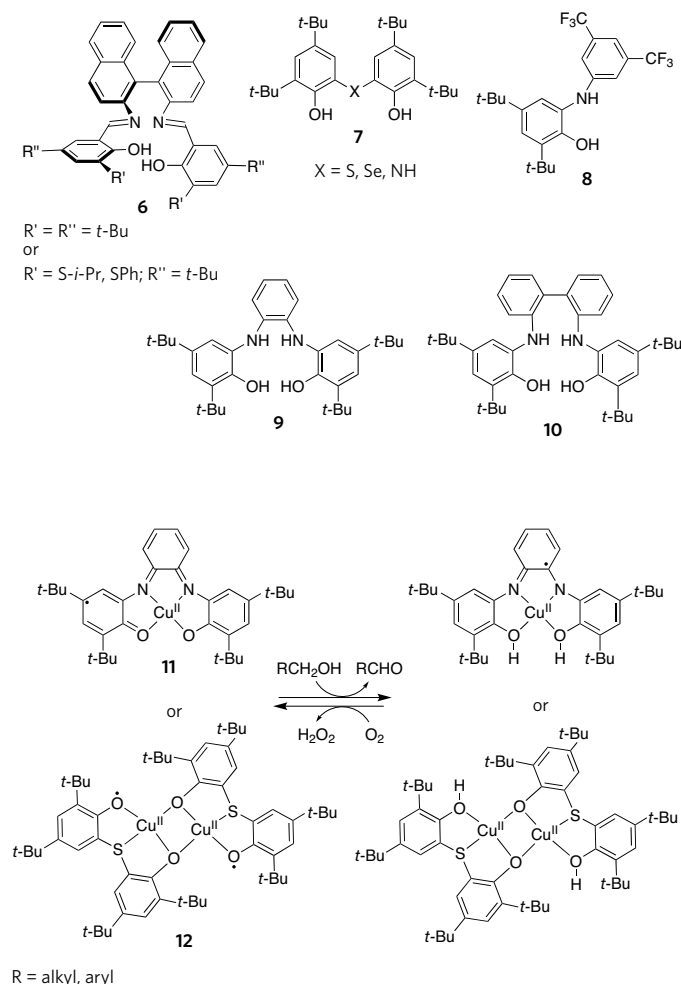


Figure 4 | Ligands for modelling GAO. Deprotonated versions of the ligands 6–10 have been used to prepare copper complexes that mimic GAO function, insofar as they catalyse the aerobic oxidation of alcohols to aldehydes or ketones via the intermediacy of Cu(II)-phenoxyl radical species and produce H₂O₂ as a coproduct. Synthetic Cu(II)-phenoxyl radical complexes of differing nuclearity (11 and 12) catalytically oxidize alcohols just as GAO does, but without the apparent involvement of a Cu(I) intermediate; instead, only ligand-centred redox reactions have been suggested to occur. *i*-Pr, isopropyl; Ph, phenyl; *t*-Bu, *tert*-butyl.

Copper enzymes

Copper-containing oxidases and oxygenases comprise a large class of enzymes that use intriguing mechanisms to bind and activate O₂ and oxidize organic substrates¹⁷. The active sites of these enzymes contain varying numbers of copper ions, and they have diverse structures that underlie similarly diverse functional attributes. We focus here on a select few for which extensive structure and function information has been obtained and applied to the development of biomimetic catalysts.

Galactose oxidase (GAO) is a well-studied member of the family of radical copper oxidases that use a novel copper(II)-tyrosyl radical unit to perform two-electron redox chemistry^{18,19}. GAO couples the reduction of O₂ to H₂O₂ with the oxidation of primary alcohols to aldehydes. As shown by X-ray crystallography (Fig. 3), the so-called 'inactive' form of GAO features an unusual tyrosinate ligand (at amino acid 272 of the enzyme (Y272)) covalently linked to a cysteine (C228)²⁰. In the 'active' form of the enzyme, this ligand exists as a one-electron oxidized tyrosyl radical that is stabilized by a nearby tryptophan²¹ (W290). In the consensus mechanism (Fig. 3), the 'active' Cu(II)–Y272• form is responsible for alcohol oxidation in one phase of the catalytic cycle. It is regenerated, with production of H₂O₂, in the second phase by reaction of the resulting

'reduced' Cu(I) form with O₂. Many details of these processes have been elucidated through studies of the enzyme, synthetic modelling^{22,23} and theoretical calculations²⁴, including dissection of the key C–H bond-activating process into electron transfer, proton transfer and H-atom abstraction components using kinetic isotope effects and quantitative structure-activity relationship correlations²⁵. Notably, the coupling of the one-electron redox cycles, Cu(I) ↔ Cu(II) and Y272 ↔ Y272•, to effect the overall two-electron catalytic half-reaction uniquely illustrates metal and organic cofactor synergism worthy of incorporation into designs of synthetic oxidation catalysts.

Tyrosinase and catechol oxidase use an oxidant entirely different from the Cu(II)-tyrosyl radical. In common with the reversible O₂ carrier protein haemocyanin, the dicopper(I) active sites of tyrosinase and catechol oxidase react with O₂ to generate a (peroxo)dicopper(II) unit that features an unusual side-on (μ-η²:η²) binding mode, as shown by X-ray crystallography for haemocyanin^{26,27} and tyrosinase²⁸ and by spectroscopy for catechol oxidase²⁹ (4, Fig. 2). The μ-η²:η²-peroxide has an O–O stretching frequency (~750 cm⁻¹) lower than that of most metal-peroxide complexes (850–900 cm⁻¹), suggesting significant activation of the dioxygen O–O bond by its side-on coordination to two Cu(II) ions. The (μ-η²:η²-peroxo)dicopper(II) unit is postulated to be responsible for the oxidation of phenols (by tyrosinase) and/or catechols (by catechol oxidase)^{17,30}, the former occurring by an electrophilic aromatic substitution pathway³¹.

Numerous (peroxo)dicopper(II) complexes have been synthesized, including examples with (μ-η²:η²-peroxo)dicopper(II) cores^{32–34}. Some of these interconvert with di(μ-oxo)dicopper(III) isomers (Fig. 1), raising the alternative possibility that this core is responsible for substrate attack during catalysis by tyrosinase or catechol oxidase (despite the fact that it has not yet been observed in a protein)³⁵. Developing synthetic systems that can access either or both cores from O₂ is an important biomimetic strategy for preparing environmentally friendly oxidizing reagents.

Monocopper–oxygen species are implicated as intermediates in the functionally important enzymes amine oxidase (AO)³⁶, dopamine β-monooxygenase (DβM) and peptidylglycine α-hydroxylating monooxygenase (PHM)³⁷. Intermediates under consideration include [CuO₂]⁺, [CuOOH]⁺ and [CuO]⁺ units (Fig. 1). Kinetic³⁸ and theoretical studies³⁹ have been interpreted to indicate that a [CuO₂]⁺ moiety attacks the substrate in PHM and DβM; such a core has been thoroughly defined in synthetic model complexes⁴⁰ and by X-ray crystallography for PHM (5, Fig. 2), although its charge is not known in this case⁴¹. Other theoretical studies suggest that the [CuO]⁺ unit is a more potent oxidant potentially capable of attacking substrate in these enzymes^{42,43}. Such a species has not been identified conclusively in an enzyme or synthetic complex, however; so its role in biological and synthetic systems remains controversial⁴⁴. Still, the anticipated high oxidizing power of the [CuO]⁺ moiety makes it an attractive target in efforts to design catalysts for the oxidation of relatively unreactive substrates.

Biologically inspired catalysis

The results of the biochemical studies discussed here provide a useful starting point for the design of synthetic oxidation catalysts. Progress in this area is summarized in this section.

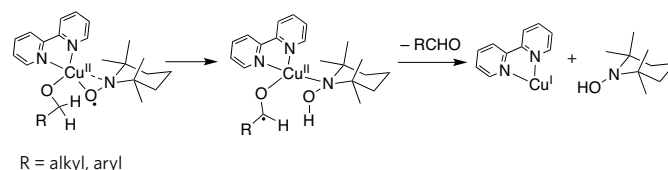


Figure 5 | Copper-TEMPO radical oxidation. The alcohol oxidation steps proposed for the bipyridine–Cu–TEMPO (Cu–2,2,6,6-tetramethyl-1-piperidinyloxy) system are closely related to those proposed for GAO (Fig. 3), despite the different chemical natures of the organic radical cofactors involved (that is, nitroxyl for Cu–TEMPO compared with phenoxyl for GAO).

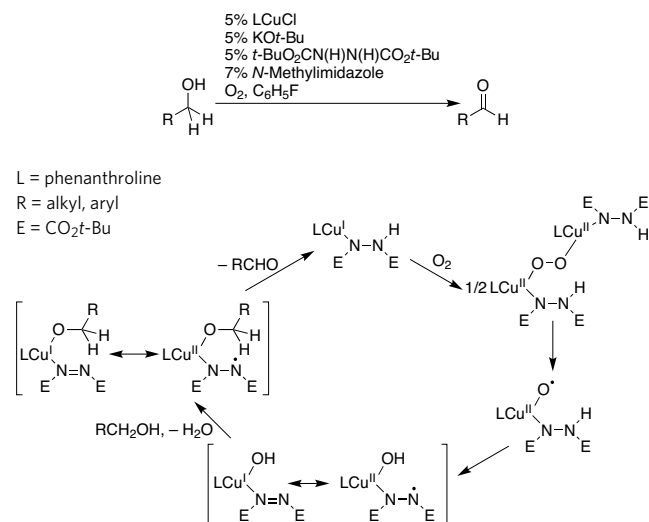


Figure 6 | Proposed mechanism for alcohol oxidation. An optimized system for the catalytic oxidation of a variety of alcohols is proposed to involve formation of a (peroxo)dicopper(II) complex that undergoes O–O bond homolysis to yield an intermediate with a [CuO]⁺ core. Intramolecular hydrogen-atom abstraction is suggested to occur, followed by ligand substitution and subsequent aldehyde generation via a mechanism conceptually related to that followed by GAO and the Cu–TEMPO systems.

Copper-catalysed alcohol oxidations

Within the broader context of copper-catalysed oxidations of hydrocarbons⁴⁵, particular attention has been paid to the oxidation of alcohols to carbonyl compounds using O₂ as an oxidant. This is because of the significance of the transformation and the desire to develop ‘green’ alternatives to procedures that use heavy metal reagents or sensitive and/or expensive oxidants, and to procedures that are not economical or environmentally friendly^{46,47}. Inspired by the coupling of an organic cofactor and Cu(I) ↔ Cu(II) redox chemistry in alcohol oxidation by the radical copper oxidases (for example, GAO), a variety of synthetic systems incorporating organic radicals and copper complexes have been studied for this reaction.

In direct analogy to the function of GAO, copper complexes of the deprotonated forms of the ligands **6–10** (Fig. 4) catalyse the aerobic oxidation of selected alcohols to aldehydes and/or ketones with concomitant generation of H₂O₂ as a coproduct^{48–54}. The involvement of Cu(II)–phenoxyl radical intermediates in these reactions is a further similarity with GAO, although differences in mechanism among the synthetic systems and the enzyme have been found. For instance, although the system derived from ligand **6** follows a pathway entirely analogous to that proposed for GAO involving cycling between Cu(II)–phenoxyl radical and Cu(I)–phenol intermediates^{48,49}, complexes **11** (ref. 53) and **12** (ref. 51) follow mechanisms that do not involve changes in the metal-oxidation state and differ further with respect to the nuclearity of the proposed active species.

Conceptually related catalytic systems for the aerobic oxidation of alcohols incorporate copper salts and the 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) radical^{46,55}. A variety of copper salts, exogenous bases and reaction media (including environmentally friendly ionic liquids⁵⁶ or fluorinated biphasic systems⁵⁷) have been used for this reaction, which differs from that of GAO insofar as H₂O rather than H₂O₂ is produced as the coproduct with the aldehyde or ketone. Nonetheless, pronounced mechanistic similarities with GAO are evident in some cases (Fig. 5), with TEMPO playing the same part as the phenoxyl radical in the enzyme.

Another versatile catalytic system combines CuCl, phenanthroline, a suitable base and a substituted hydrazine as the reducing agent⁵⁸. The optimized conditions shown in Fig. 6 include *N*-methylimidazole as an additive and are useful for the clean oxidation of a variety of primary and

secondary alcohols to the respective ketones or aldehydes⁵⁹. In common with the Cu–TEMPO systems, O₂ acts as stoichiometric oxidant and H₂O is the sole byproduct. An intriguing mechanism has been proposed involving a [CuO]⁺ species as an intermediate and a pathway for intramolecular oxidation of coordinated alkoxide that bears some resemblance to that proposed for GAO and the Cu–TEMPO systems.

Copper-catalysed phenol and catechol oxidations

The oxidative reactivity of dicopper complexes that model the active sites of tyrosinase and catechol oxidase has been explored extensively in efforts aimed mostly at obtaining mechanistic insights into the activity of the enzymes^{33,34,60,61}. With respect to understanding the phenolase activity of tyrosinase, emphasis has been placed on examining the involvement of the (μ-η²:η²-peroxo)dicopper(II) complexes and/or their di(μ-oxo)dicopper isomers in stoichiometric (non-catalytic) hydroxylations of ligand arene moieties and external phenols. In a catalytic application that capitalizes on the common observation of radical coupling of phenol substrates by these oxidants, tyrosinase model complexes have been used for the regio-controlled oxidative coupling polymerization of phenols^{62–66} (Fig. 7a). Steric influences of the supporting ligand are thought to be critical in inhibiting undesired reactions at the *ortho* position in the putative Cu(II)–phenolate ↔ Cu(I)–phenoxyl intermediate, thus resulting in useful poly(1,4-phenylene ether) materials. Several dicopper complexes have also been studied as catalysts for the aerobic oxidation of catechols to *ortho*-quinones, which is analogous to catechol oxidase activity. In general, catechols may be oxidized both by dicopper(II) complexes and O₂ adducts such as (μ-peroxo)dicopper(II) species, and catalytic cycles incorporating both processes have been suggested that typically involve coordination of the catecholate to both copper ions and evolution of H₂O and/or H₂O₂ byproducts (Fig. 7b).

Iron-catalysed hydrocarbon oxidations

A common feature of the catalytic cycles of cytochrome P450, MMO and the Rieske dioxygenases is the involvement of an iron(III)–peroxo intermediate (Fig. 1). This intermediate may react directly with the substrate

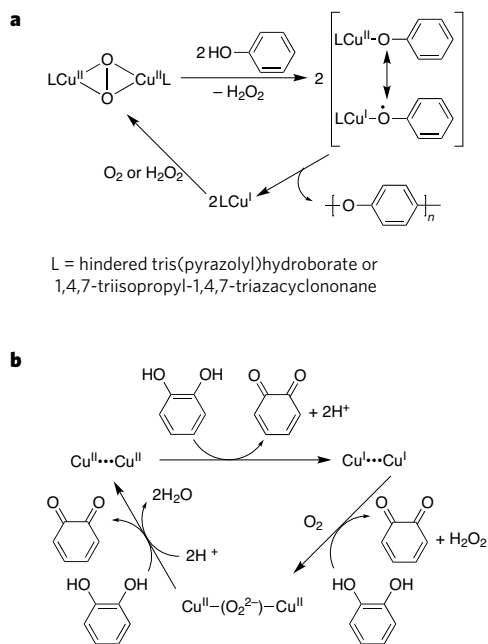


Figure 7 | Oxidations by dicopper complexes. **a**, The proposed catalytic cycle for the regio-controlled oxidative polymerization of phenols involves a (μ-η²:η²-peroxo)dicopper(II) model of the oxy forms of tyrosinase, haemocyanin and catechol oxidase. **b**, Multiple oxidation pathways have been proposed for the catalytic oxidation of catechols to *ortho*-quinones by dicopper complexes⁶⁰.

Table 1 | Iron and manganese oxidation catalysts that use H₂O₂ as the oxidant to give high conversion of alkenes into epoxide or *cis*-diol products

Catalyst (mol%)	Additive (equivalents per metal)	Solvent (Temperature)	Alkene	H ₂ O ₂ :alkene	Epoxide yield	<i>cis</i> -Diol yield	Reference
Fe(13) (0.05)	None	3:1 MeOH:MeCN (25 °C)	Cyclooctene	1.2	95%	0%	71
Fe(14) (1)	None	CH ₂ Cl ₂ * (25 °C)	Cyclooctene	1	81%	0%	72
Fe(17) (3)	None	MeCN (25 °C)	1-Octene	1.5	73%	3%	84
Fe(17) (3)	HOAc (10)	MeCN (25 °C)	1-Decene	1.5	85%	0%	83
Fe(17) (0.5)	HOAc (12,000)	MeCN (0 °C)	Cyclooctene	1.5	99%	<1%	85
Fe(15) (3)	None	MeCN (25 °C)	1-Octene	4	16%	53%	83
Fe(15) (0.5)	HOAc (12,000)	MeCN (0 °C)	Cyclooctene	1.5	99%	<1%	85
Fe(16) (5)	None	MeCN (25 °C)	Cyclooctene	1.5	9%	0%	87
Fe(16) (5)	HOTf (5)	MeCN (25 °C)	Cyclooctene	1.5	86%	0%	87
1:2 FeCl ₃ : 19 (5)	Pyrrolidine (2)	<i>t</i> -AmylOH (25 °C)	<i>trans</i> -Stilbene	2	97%	0%	89
MnSO ₄ (1)	0.2 M HCO ₃ [−] pH 8 buffer (0.25)	DMF (25 °C)	Cyclohexene	10	99%	0%	92
Mn(20) (0.1)	None	MeCN (0 °C)	Cyclooctene	1.3	1%	0.5%	93
Mn(20) (0.1)	Cl ₃ CCOOH (10)	MeCN (0 °C)	Cyclooctene	1.3	24%	44%	94
Mn(20) (0.1)	2,6-Cl ₂ C ₆ H ₃ -COOH (30)	MeCN (0 °C)	Cyclooctene	1.3	8%	53%	94
Mn(20) (0.1)	Salicylic acid (50)	MeCN (0 °C)	Cyclooctene	1.3	51%	13%	94

*In a biphasic mixture with the ionic liquid 1-butyl-3-methylimidazolium bromide. DMF, dimethylformamide; Me, methyl; OAc, acetate; OTf, trifluoromethanesulphonic acid.

or, more likely, undergo O–O bond cleavage to generate a high-valent iron–oxo species that contains the two oxidizing equivalents required to oxidize the substrate, the locations of which differ for each enzyme. The two oxidizing equivalents are distributed between the iron centre and the porphyrin ligand for cytochrome P450 and between the two iron centres in MMO, but must be localized on the mononuclear iron centre of the Rieske dioxygenases. There are thus several possible strategies for developing biomimetic catalysts. In this section, we focus on recent developments in iron catalysis using H₂O₂ as an oxidant. From a biochemical perspective, these biologically inspired catalytic systems correspond to the peroxide shunt pathways that have been demonstrated for all three enzyme types and involve reactions of the iron(III) forms^{7,8,67,68} of the enzymes with H₂O₂. The use of a combination of an iron salt or complex and H₂O₂ as an oxidizing system has a history that dates back to the late nineteenth century⁶⁹, but the major challenge has been to inhibit the homolytic cleavage of the peroxo O–O bond that produces non-selective, and thus unwanted, hydroxyl radicals. Instead, the aim is to direct the metal-promoted cleavage towards the generation of a metal-based oxidant that can carry out hydrocarbon oxidations with high chemo-, regio- and stereoselectivity^{7,8}.

Much effort has been invested in the development of metalloporphyrin catalysts that mimic the reactivity of cytochrome P450, and these endeavours have been reviewed extensively⁷⁰. The first two examples in Table 1 illustrate how efficiently H₂O₂ can be converted into desired epoxide products by using electron-deficient porphyrins^{71,72} (ligands are shown in Fig. 8). For the catalyst Fe(**13**), the solvent methanol provides the protons needed to promote the heterolytic cleavage of the peroxo O–O bond^{71,73}, thereby mimicking a mechanistic principle well established for cytochrome P450 and related haem peroxidases^{7,74}.

The marked increase in information on non-haem iron oxygenases within the past decade^{10,75,76} has spurred efforts to explore the use of non-haem ligand scaffolds to facilitate such oxidation catalysis. Only recently have conditions been found to engender oxidative transformations with high stereoselectivity^{77–79}. In this section, we highlight these notable successes in the epoxidation and *cis*-dihydroxylation of C=C bonds and the hydroxylation of aliphatic C–H bonds, which are all potentially important transformations in organic synthesis and medicinal chemistry.

The most extensively studied complexes thus far are those of tetradentate nitrogen-donor ligands with topologies that allow two *cis*-oriented coordination sites to be available for peroxide binding and activation. This arrangement is analogous to that found for the Rieske dioxygenases as illustrated by **3** (Fig. 2) but is in contrast to that typical for haem enzymes where only one coordination site is available for this purpose. A key observation that led to the discovery of this family of complexes

was the use of dilute H₂O₂, which minimized side reactions such as the unwanted production of highly reactive hydroxyl radicals⁸⁰. When the reaction was carried out in this manner, compelling evidence for a metal-based oxidant was obtained. Specifically for the prototypical Fe(**15**) catalyst, cyclohexane oxidation resulted in a high alcohol:ketone product ratio that was unaffected by the presence of O₂, and hydroxylation of the tertiary C–H bonds of *cis*-1,2-dimethylcyclohexane

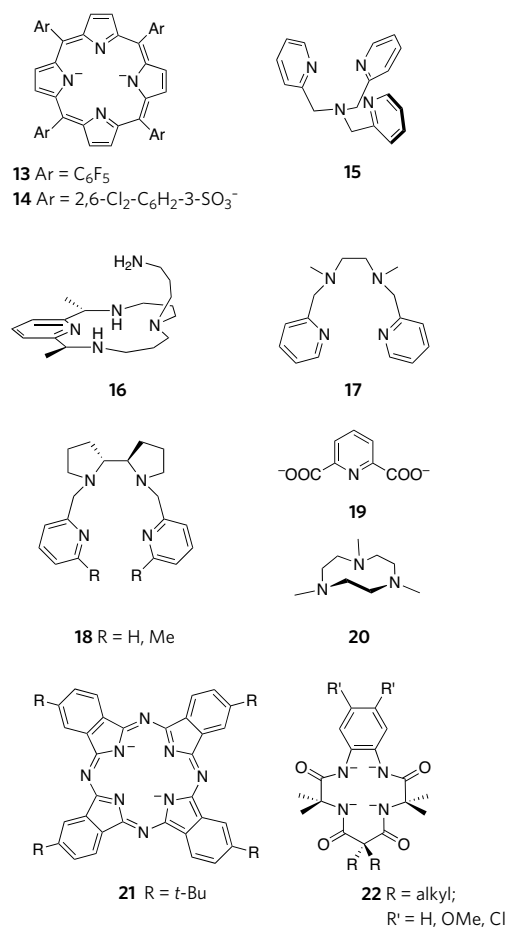


Figure 8 | Ligands used for iron- and manganese-catalysed oxidations. Many ligands, 13–22, have been used to support iron or manganese centres that catalyse the oxidation of hydrocarbons by H₂O₂ (Table 1).

afforded a tertiary alcohol product with retention of configuration of the *cis*-1,2-dimethyl groups⁸¹. These results indicate a C–H bond cleavage event where alkyl radicals (if formed) are very short-lived. Oxidation of alkenes yielded epoxide and *cis*-diol products with similarly high retention of configuration⁸². Importantly, these reactions are the first examples of iron-catalysed *cis*-dihydroxylation of an alkene.

The nature of the metal-based oxidant in Fe(15) catalysis was deduced from a combination of low-temperature spectroscopic studies and room-temperature ¹⁸O-labelling experiments^{81,82}. At –40 °C an Fe(III)–OOH intermediate was trapped in acetonitrile (MeCN) solvent and characterized by a variety of spectroscopic methods; this intermediate was then proposed to undergo O–O bond cleavage to form the putative Fe(v)(O)(OH) oxidant. Strong support for the involvement of the latter came from labelling experiments that established the incorporation of added H₂¹⁸O into the oxidation products with retention of stereochemistry (Fig. 9). Most compelling was the observation that the *cis*-diol product incorporated one oxygen from H₂O₂ and the other from H₂O. These results emphasize the critical part a proximal proton donor (that is, the metal-bound water) plays in promoting the heterolytic cleavage of the O–O bond and rationalizes why complexes of closely related pentadentate ligands are poor catalysts for such oxidations.

So far, several non-haem iron-based catalytic oxidation systems have been developed with potential practical importance. Eric Jacobsen and co-workers demonstrated the use of acetic acid as an additive (30 mol%) in Fe(II)(17) catalysis that converted alkenes into epoxides with a high yield⁸³. With 3 mol% catalyst and a 50% excess of H₂O₂ at 4 °C, even terminal alkenes could be epoxidized to give product yields as high as 90%. The added acetic acid is clearly important, as this additive enhances both the yield and selectivity for epoxide^{83,84}.

Further exploration of the effect of acetic acid in Fe(17) and Fe(15) catalysis by Rubén Mas-Ballesté and Lawrence Que Jr.⁸⁵ showed that cyclooctene could be converted nearly quantitatively to an epoxide within 2 min in a 1:2 MeCN:acetic acid solvent mixture at 0 °C. Spectroscopic and kinetic evidence was obtained for the binding of acetic acid to the Fe(III)–OOH intermediate, emphasizing the importance of a proximal proton donor to promote the heterolytic cleavage of the O–O bond to form an Fe(v)(O)(OAc) oxidant^{85,86} (Fig. 9) and reiterating a principle well documented from studies of cytochrome P450 and haem peroxidases^{77,74}. Similarly, Elena Rybak-Akimova and co-workers showed that the iron complex of macrocycle 16 (Fig. 8) became an effective epoxidation catalyst only when five equivalents of trifluoromethanesulphonic acid were added to the reaction mixture⁸⁷. The added acid was proposed to protonate the aminoalkyl ‘tail’, generating an ammonium group close enough to a presumed Fe–OOH intermediate to promote heterolytic O–O bond cleavage.

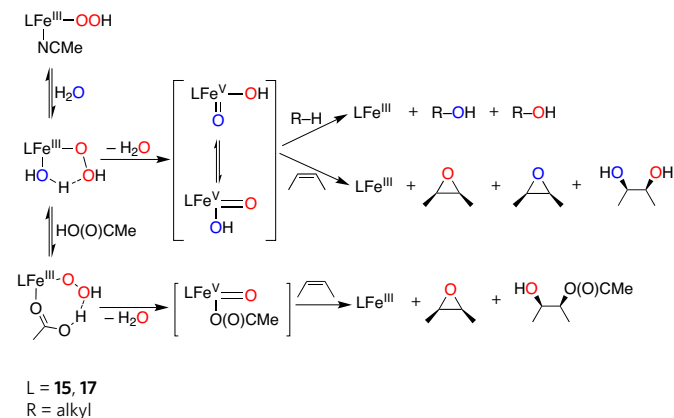


Figure 9 | Mechanisms of oxidations by iron catalysts. Proposed mechanisms for stereospecific oxidations of alkanes and alkenes by non-haem iron complexes invoke Fe(v)=O oxidants. Particularly instructive is the formation of a *cis*-diol from an alkene, with one oxygen atom from H₂O₂ and the other from H₂O. Oxygen atoms derived from H₂O₂ (not depicted) are shown in red, and oxygen atoms derived from H₂O are shown in blue. Me, methyl.

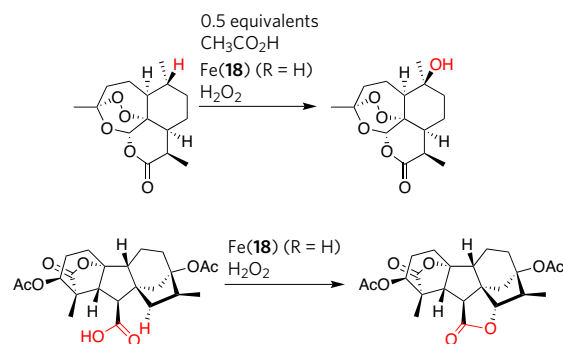


Figure 10 | Illustrative selective oxidations by iron. Highly selective oxidations of secondary and tertiary C–H bonds in complex natural products have been achieved using Fe(18) (R = H) and H₂O₂, as illustrated by the two reactions shown. The key functional groups involved are shown in red. Ac, acetyl.

Mark Chen and M. Christina White⁸⁸ have ingeniously applied the above Fe(17)–H₂O₂–acetic acid chemistry⁴⁶ to catalyse the hydroxylation of unactivated tertiary C–H bonds with a related Fe(18, where R is H) complex. With 5 mol% catalyst, the target C–H bond could be hydroxylated in 40–60% yield in many of the examples presented. In general, the most susceptible tertiary C–H bond is the one that is most electron-rich and least sterically hindered. The transformation was tolerant of many functional groups, as shown in Fig. 10 (upper panel) for the hydroxylation of the antimalarial drug artemisinin at the C10 position. When the substrate had a carboxylate functionality, the carboxylate took the place of the acetic acid additive and could be used to directly attack a nearby C–H bond, as illustrated in Fig. 10 (lower panel) for the hydroxylation of a tetrahydrogibberellic acid analogue. These examples show that these biologically inspired catalysts may be useful for the synthesis of complex organic molecules.

Iron complexes can also be used to catalyse asymmetric alkene oxidation, and two systems affording high enantioselectivity are mentioned here. Matthias Beller and co-workers developed an *in situ*-generated catalyst capable of epoxidizing aryl alkenes with up to 97% yield (Table 1) by using a combination of FeCl₃, two equivalents of pyridine-2,6-carboxylic acid (the doubly protonated form of 19) and two equivalents of pyrrolidine in *tert*-amyl alcohol solvent⁸⁹. When monotosylated (*S,S*)-1,2-diphenyldiaminoethane was used in place of pyrrolidine⁹⁰, epoxides with an enantiomeric excess as high as 97% were obtained. By contrast, Que and co-workers used chiral ligand 18 (where R is Me) (Fig. 8) to obtain a complex that catalysed the oxidation of *trans*-2-heptene to give a *cis*-diol product with 97% enantiomeric excess⁹¹. Insights into the nature of the active species in these systems are not yet available.

Although not a strategy often used in nature, substituting Mn for Fe can result in superior oxidation catalysts, as was found for metalloporphyrins⁷⁰. Non-porphyrinic Mn centres can also activate H₂O₂ to catalyse alkene oxidations. Kevin Burgess and co-workers found that even simple MnSO₄ was effective for alkene epoxidation, when carried out in dimethylformamide (DMF) solvent in the presence of 0.2 M NaHCO₃ buffer at pH 8 (ref. 92). HCO₄[–] was formed under these conditions and presumably reacted with the Mn(II) ion to form either a Mn(II)–peroxocarbonato or Mn(IV)=O species that carried out the epoxidation. Ben Feringa and co-workers demonstrated that dinuclear Mn(20) complexes catalysed alkene oxidations with better turnover numbers than Fe(15) and Fe(17) (700 compared with 200 turnovers in the best-case scenarios)^{93,94}. Both epoxide and *cis*-diol products were obtained, and the catalytic efficiency and the epoxide:diol ratio were modulated by the electronic and steric nature of carboxylic acid additives. Labelling studies of cyclooctene oxidation showed that H₂¹⁸O incorporation into epoxide and *cis*-diol products analogous to those reported for Fe(15), implicating the involvement of metal–peroxo and high-valent metal–oxo intermediates⁸². However, unlike in the iron study, the various species involved in Mn catalysis were proposed to be dinuclear in nature, and no intermediate has been identified so far⁹⁴.

Very recently, it was reported that methane could be oxidized in water by H_2O_2 at 25–60 °C with silica-supported $[\text{Fe}_2\text{N}(\mathbf{21})_2]$ as the catalyst⁹⁵. As many as 150 mols CH_4 per mol catalyst could be oxidized to formaldehyde and formic acid. Catalytic activity was improved by the introduction of 0.1 M H_2SO_4 and evidence was obtained for an H_2O_2 -derived oxo adduct of the starting catalyst by electrospray mass spectrometry. On this basis, a mechanism following the haem paradigm (Fig. 1) was proposed involving formation of metal–peroxo and high-valent metal–oxo species on the intact Fe–N–Fe unit.

In closing, we briefly mention here the use of iron complex and H_2O_2 combinations for environmental applications, which go beyond the scope of this review. This effort has been pioneered by Terrence Collins, whose group has designed and developed complexes of tetraamido macrocyclic ligands (TAML) such as **22** that are water-soluble and oxidatively and hydrolytically robust⁹⁶. These complexes catalyse the activation of H_2O_2 and are being investigated for use in the treatment of waste water from pulp and textile mills to remove coloured effluents and toxic chlorinated phenols. Although the nature of the oxidant(s) generated by the Fe(**22**) and H_2O_2 combination that carry out these interesting transformations has not been established, metastable oxoiron(IV) and oxoiron(V) complexes of the TAML family have recently been trapped in organic solvents at low temperature and characterized spectroscopically^{97,98}. Indeed, the latter is the only bona fide oxoiron(V) complex that has been identified.

Challenges

In the past decade, Fe, Mn and Cu complexes with non-porphyrinic ligands that can catalyse oxidations of C–H and C=C bonds by O_2 or H_2O_2 have been identified. These oxidations were inspired by a mechanistic understanding of similar transformations carried out by iron and copper enzymes. The successes noted in this review are merely starting points for further investigations; much more work needs to be done to accomplish the goal of discovering broadly useful, active and highly selective hydrocarbon oxidation catalysts that operate in an environmentally friendly manner. With H_2O_2 as the oxidant, appropriate conditions have been determined that minimize formation of unselective hydroxyl radicals and promote the heterolytic cleavage of the O–O bond. The latter generates a more selective metal-based oxidant that can effect substrate oxidations with significant chemoselectivity, regioselectivity, stereoselectivity and/or enantioselectivity. Putting this into practice in ways that are useful for industrial and fine-chemical synthesis is an important goal. Similarly, although effective catalysts for alcohol oxidations using copper–cofactor pairs have been developed, finding systems that are both highly active and broadly applicable requires further effort⁴⁶. More generally, extending the range of substrates that can be oxidized under environmentally friendly conditions by iron or copper catalysts is an important goal, a prime example being the selective hydroxylation of methane to methanol performed by the soluble (di-iron-containing) and particulate (copper-containing) MMOs^{10,99}.

A significant challenge is the development of selective catalytic systems that use O_2 as the oxidant and avoid deleterious side reactions. Biological systems have unique capabilities in this regard through the control of the spatial and/or temporal distribution of substrates and oxidants. Such control has yet to be exerted in model systems and needs to be addressed, perhaps through more sophisticated biomimetic design strategies than those used so far. Further discoveries are likely to be made as mechanistic understanding of the enzymes accrues, new types of reactive intermediate in the biological and synthetic models systems are uncovered and applications toward catalysis are explored. ■

- Arakawa, H. *et al.* Catalysis research of relevance to carbon management: progress, challenges, and opportunities. *Chem. Rev.* **101**, 953–996 (2001).
- Punniyamurthy, T., Velusamy, S. & Iqbal, J. Recent advances in transition metal catalyzed oxidation of organic substrates with molecular oxygen. *Chem. Rev.* **105**, 2329–2364 (2005).
- Conley, B. L. *et al.* In *Activation of Small Molecules: Organometallic and Bioinorganic Perspectives* (ed. Tolman, W. B.) 235–285 (Wiley-VCH, 2006).
- Meunier, B. (ed.) *Biomimetic Oxidations Catalyzed by Transition Metal Complexes* (Imperial College Press, 2000).
- Mahadevan, V., Klein Gebbink, R. J. M. & Stack, T. D. P. Biomimetic modeling of copper oxidase reactivity. *Curr. Opin. Chem. Biol.* **4**, 228–234 (2000).
- Sheldon, R. A., Arends, I. & Hanefeld, U. *Green Chemistry and Catalysis* (Wiley-VCH, 2007).
- Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* **105**, 2253–2278 (2005).
- Dawson, J. H. Probing structure–function relations in heme-containing oxygenases and peroxidases. *Science* **240**, 433–439 (1988).
- Que, L. The heme paradigm revisited: alternative reaction pathways considered. *J. Biol. Inorg. Chem.* **9**, 643–690 (2004).
- Merkx, M. *et al.* Dioxxygen activation and methane hydroxylation by soluble methane monooxygenase: a tale of two irons and three proteins. *Angew. Chem. Int. Edn Engl.* **40**, 2782–2807 (2001).
- Beauvais, L. G. & Lippard, S. J. Reactions of the peroxo intermediate of soluble methane monooxygenase hydroxylase with ethers. *J. Am. Chem. Soc.* **127**, 7370–7378 (2005).
- Waller, B. J. & Lipscomb, J. D. Dioxxygen activation by enzymes containing binuclear non-heme iron clusters. *Chem. Rev.* **96**, 2625–2658 (1996).
- Shu, L. *et al.* An $\text{Fe}_2^{IV}\text{O}_2$ diamond core structure for the key intermediate Q of methane monooxygenase. *Science* **275**, 515–518 (1997).
- Ferraro, D. J., Gakhar, L. & Ramaswamy, S. Rieske business: structure–function of Rieske non-heme oxygenases. *Biochem. Biophys. Res. Commun.* **338**, 175–190 (2005).
- Karlsson, A. *et al.* Side-on binding of dioxxygen to iron: implications for enzymatic cis-dihydroxylation reactions. *Science* **299**, 1039–1042 (2003).
- Krebs, C., Galonic Fujimori, D., Walsh, C. T. & Bollinger, J. M. Jr. Non-heme Fe(IV)–oxo intermediates. *Acc. Chem. Res.* **40**, 484–492 (2007).
- Solomon, E. I., Chen, P., Metz, M., Lee, S.-K. & Palmer, A. E. Oxygen binding, activation, and reduction to water by copper proteins. *Angew. Chem. Int. Edn Engl.* **40**, 4570–4590 (2001).
- Whittaker, J. W. Free radical catalysis by galactose oxidase. *Chem. Rev.* **103**, 2347–2363 (2003).
- Whittaker, J. W. The radical chemistry of galactose oxidase. *Arch. Biochem. Biophys.* **433**, 227–239 (2005).
- Ito, N. *et al.* Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. *Nature* **350**, 87–90 (1991).
- Rogers, M. S. *et al.* The stacking tryptophan of galactose oxidase: a second-coordination sphere residue that has profound effects on tyrosyl radical behavior and enzyme catalysis. *Biochemistry* **46**, 4606–4618 (2007).
- Jazdzewski, B. A. & Tolman, W. B. Understanding the copper-phenoxyl radical array in galactose oxidase: contributions from synthetic modeling studies. *Coord. Chem. Rev.* **200–202**, 633–685 (2000).
- Fabrice, T. Ten years of a biomimetic approach to the copper(II) radical site of galactose oxidase. *Eur. J. Inorg. Chem.* **2007**, 2379–2404 (2007).
- Rokhsana, D., Dooley, D. M. & Szilagyi, R. K. Systematic development of computational models for the catalytic site in galactose oxidase: impact of outer-sphere residues on the geometric and electronic structures. *J. Biol. Inorg. Chem.* **13**, 371–383 (2008).
- Whittaker, M. M. & Whittaker, J. W. Catalytic reaction profile for alcohol oxidation by galactose oxidase. *Biochemistry* **40**, 7140–7148 (2001).
- Magnus, K. A., Ton-That, H. & Carpenter, J. E. Recent structural work on the oxygen transport protein hemocyanin. *Chem. Rev.* **94**, 727–735 (1994).
- Cuff, M. E., Miller, K. I., van Holde, K. E. & Hendrickson, W. A. Crystal structure of a functional unit from octopus hemocyanin. *J. Mol. Biol.* **278**, 855–870 (1998).
- Matoba, Y., Kumagai, T., Yamamoto, A., Yoshitsu, H. & Sugiyama, M. Crystallographic evidence that the dinuclear copper center of tyrosinase is flexible during catalysis. *J. Biol. Chem.* **281**, 8981–8990 (2006).
- Rompel, A. *et al.* Purification and spectroscopic studies on catechol oxidases from *Lycopus europaeus* and *Populus nigra*: evidence for a dinuclear copper center of type 3 and spectroscopic similarities of tyrosinase and hemocyanin. *J. Biol. Inorg. Chem.* **4**, 56–63 (1999).
- Granata, A., Monzani, E., Bubacco, L. & Casella, L. Mechanistic insight into the activity of tyrosinase from variable-temperature studies in an aqueous/organic solvent. *Chem. Eur. J.* **12**, 2504–2514 (2006).
- Yamazaki, S. & Itoh, S. Kinetic evaluation of phenolase activity of tyrosinase using simplified catalytic reaction system. *J. Am. Chem. Soc.* **125**, 13034–13035 (2003).
- Mirica, L. M., Ottenwaelde, X. & Stack, T. D. P. Structure and spectroscopy of copper-dioxxygen complexes. *Chem. Rev.* **104**, 1013–1045 (2004).
- Lewis, E. A. & Tolman, W. B. Reactivity of copper-dioxxygen systems. *Chem. Rev.* **104**, 1047–1076 (2004).
- Hatcher, L. Q. & Karlin, K. D. Ligand influences in copper-dioxxygen complex-formation and substrate oxidations. *Adv. Inorg. Chem.* **58**, 131–184 (2006).
- Mirica, L. M. *et al.* Tyrosinase reactivity in a model complex: an alternative hydroxylation mechanism. *Science* **308**, 1890–1892 (2005).
- Brazeau, B. J., Johnson, B. J. & Wilmot, C. M. Copper-containing amine oxidases. Biogenesis and catalysis: a structural perspective. *Arch. Biochem. Biophys.* **428**, 22–31 (2004).
- Klinman, J. P. The copper-enzyme family of dopamine β-monooxygenase and peptidylglycine α-hydroxylating monooxygenase: resolving the chemical pathway for substrate hydroxylation. *J. Biol. Chem.* **281**, 3013–3016 (2006).
- Evans, J. P., Ahn, K. & Klinman, J. P. Evidence that dioxygen and substrate activation are tightly coupled in dopamine β-monooxygenase: implications for the reactive oxygen species. *J. Biol. Chem.* **278**, 49691–49698 (2003).
- Chen, P. & Solomon, E. I. Oxygen activation by the noncoupled binuclear copper site in peptidylglycine α-hydroxylating monooxygenase. Reaction mechanism and role of the noncoupled nature of the active site. *J. Am. Chem. Soc.* **126**, 4991–5000 (2004).
- Cramer, C. J. & Tolman, W. B. Mononuclear CuO_2 complexes: geometries, spectroscopic properties, electronic structures, and reactivity. *Acc. Chem. Res.* **40**, 601–608 (2007).
- Prigge, S. T., Eipper, B. A., Mains, R. E. & Amzel, L. M. Dioxxygen binds end-on to mononuclear copper in a precatalytic enzyme complex. *Science* **304**, 864–867 (2004).
- Yoshizawa, K., Kihara, N., Kamachi, T. & Shiota, Y. Catalytic mechanism of dopamine β-monooxygenase mediated by Cu(III) -oxo. *Inorg. Chem.* **45**, 3034–3041 (2006).

43. Crespo, A., Marti, M. A., Roitberg, A. E., Amzel, L. M. & Estrin, D. A. The catalytic mechanism of peptidylglycine α -hydroxylating monooxygenase investigated by computer simulation. *J. Am. Chem. Soc.* **128**, 12817–12828 (2006).
44. Rolff, M. & Tuzcek, F. How do copper enzymes hydroxylate aliphatic substrates? Recent insights from the chemistry of model systems. *Angew. Chem. Int. Edn Engl.* **47**, 2344–2347 (2008).
45. Punniyamurthy, T. & Rout, L. Recent advances in copper-catalyzed oxidation of organic compounds. *Coord. Chem. Rev.* **252**, 134–154 (2008).
46. Arends, I., Gamez, P. & Sheldon, R. A. Green oxidation of alcohols using biomimetic Cu complexes and Cu enzymes as catalysts. *Adv. Inorg. Chem.* **58**, 235–279 (2006).
This review surveys progress in the use of biologically inspired copper-ligand radical complexes as alcohol oxidation catalysts.
47. Piera, J. & Bäckvall, J. E. Catalytic oxidation of organic substrates by molecular oxygen and hydrogen peroxide by multistep electron transfer — a biomimetic approach. *Angew. Chem. Int. Edn Engl.* **47**, 3506–3523 (2008).
This topical review on biologically inspired oxidation catalysis emphasizes systems that use electron-transfer mediators to couple reduction and oxidation steps in biomimetic catalytic cycles.
48. Wang, Y. & Stack, T. D. P. Galactose oxidase model complexes: catalytic reactivities. *J. Am. Chem. Soc.* **118**, 13097–13098 (1996).
49. Wang, Y., DuBois, J. L., Hedman, B., Hodgson, K. O. & Stack, T. D. P. Catalytic galactose oxidase models: biomimetic Cu(II)-phenoxyl-radical reactivity. *Science* **279**, 537–540 (1998).
50. Chaudhuri, P. *et al.* Biomimetic metal-radical reactivity: aerial oxidation of alcohols, amines, aminophenols and catechols catalyzed by transition metal complexes. *J. Biol. Chem.* **386**, 1023–1033 (2005).
51. Chaudhuri, P., Hess, H., Flörke, U. & Wiegardt, K. From structural models of galactose oxidase to homogeneous catalysis: efficient aerobic oxidation of alcohols. *Angew. Chem. Int. Edn Engl.* **37**, 2217–2220 (1998).
52. Chaudhuri, P., Hess, M., Weyhermüller, T. & Wiegardt, K. Aerobic oxidation of primary alcohols by a new mononuclear Cu(II)-radical catalyst. *Angew. Chem. Int. Edn Engl.* **38**, 1095–1098 (1999).
53. Chaudhuri, P. *et al.* Aerobic oxidation of primary alcohols (including methanol) by copper(II)- and zinc(II)-phenoxyl radical catalysts. *J. Am. Chem. Soc.* **121**, 9599–9610 (1999).
54. Paine, T. K., Weyhermüller, T., Wiegardt, K. & Chaudhuri, P. Aerial oxidation of primary alcohols and amines catalyzed by Cu(II) complexes of 2,2'-selenobis(4,6-di-tert-butylphenol) providing [O,Se,O]-donor atoms. *Dalton Trans.* 2092–2101 (2004).
55. Sheldon, R. & Arends, I. Catalytic oxidations mediated by metal ions and nitroxyl radicals. *J. Mol. Catal. A* **251**, 200–214 (2006).
56. Jiang, N. & Ragauskas, A. J. Copper(II)-catalyzed aerobic oxidation of primary alcohols to aldehydes in ionic liquid [bmpp]PF₆. *Org. Lett.* **7**, 3689–3692 (2005).
57. Ragagnin, G., Betzemeier, B., Quici, S. & Knochel, P. Copper-catalysed aerobic oxidation of alcohols using fluorous biphasic catalysis. *Tetrahedron* **58**, 3985–3991 (2002).
58. Markó, I. E. *et al.* Efficient, ecologically benign, aerobic oxidation of alcohols. *Adv. Inorg. Chem.* **56**, 211–240 (2004).
59. Markó, I. *et al.* Efficient, copper-catalyzed, aerobic oxidation of primary alcohols. *Angew. Chem. Int. Edn Engl.* **43**, 1588–1591 (2004).
60. Battaini, G., Granata, A., Monzani, E., Gullotti, M. & Casella, L. Biomimetic oxidations by dinuclear and trinuclear copper complexes. *Adv. Inorg. Chem.* **58**, 185–233 (2006).
61. Koval, I. A., Gamez, P., Belle, C., Selmeçci, K. & Reedijk, J. Synthetic models of the active site of catechol oxidase: mechanistic studies. *Chem. Soc. Rev.* **35**, 814–840 (2006).
References 60 and 61 review the catalytic oxidations of phenols and catechols by synthetic models of the dicopper active sites of tyrosinase and catechol oxidase from a mechanistic perspective.
62. Higashimura, H. *et al.* Highly regioselective oxidative polymerization of 4-phenoxyphenol to poly(1,4-phenylene oxide) catalyzed by tyrosinase model complexes. *J. Am. Chem. Soc.* **120**, 8529–8530 (1998).
63. Higashimura, H. *et al.* 'Radical-controlled' oxidative polymerization of 4-phenoxyphenol by a tyrosinase model complex catalyst to poly(1,4-phenylene oxide). *Macromolecules* **33**, 1986–1995 (2000).
64. Higashimura, H. *et al.* 'Radical-controlled' oxidative polymerization of *o*-cresol catalyzed by μ - η^2 - η^2 -peroxo dicopper(II) complex. *Appl. Catal. A* **194**, 427–433 (2000).
65. Higashimura, H., Fujisawa, K., Kubota, M. & Kobayashi, S. 'Radical-controlled' oxidative polymerization of phenol: Comparison with that of 4-phenoxyphenol. *J. Polym. Sci. Polym. Chem.* **43**, 1955–1962 (2005).
66. Shibasaki, Y., Suzuki, Y. & Ueda, M. Copper-catalyzed regio-controlled oxidative coupling polymerization of 2,5-dimethylphenol. *Macromolecules* **40**, 5322–5325 (2007).
67. Andersson, K. K., Froland, W. A., Lee, S.-K. & Lipscomb, J. D. Dioxxygen independent oxygenation of hydrocarbons by methane monooxygenase hydroxylase component. *New J. Chem.* **15**, 411–415 (1991).
68. Wolfe, M. D. & Lipscomb, J. D. Hydrogen peroxide-coupled *cis*-diol formation catalyzed by naphthalene 1,2-dioxygenase. *J. Biol. Chem.* **278**, 829–835 (2003).
69. Fenton, H. J. H. Oxidation of tartaric acid in the presence of iron. *J. Chem. Soc.* **65**, 889–910 (1894).
70. Meunier, B., Robert, A., Pratiel, G. & Bernadou, J. In *The Porphyrin Handbook* (eds Kadish, K. M., Smith, K. M. & Guillard, R.) 119–187 (Academic, 2000).
71. Nam, W. *et al.* Factors affecting the catalytic epoxidation of olefins by iron porphyrin complexes and H₂O₂ in protic solvents. *J. Org. Chem.* **68**, 7903–7906 (2007).
72. Srinivas, K. A., Kumar, A. & Chauhan, M. S. Epoxidation of alkenes with hydrogen peroxide catalyzed by iron(III) porphyrins in ionic liquids. *Chem. Commun.* 2456–2457 (2002).
73. Stephenson, N. A. & Bell, A. T. A study of the mechanism and kinetics of cyclooctene epoxidation catalyzed by iron(III) tetrakis(pentafluorophenyl) porphyrin. *J. Am. Chem. Soc.* **127**, 8635–8643 (2005).
74. Sono, M., Roach, M. P., Coulter, E. D. & Dawson, J. H. Heme-containing oxygenases. *Chem. Rev.* **96**, 2841–2887 (1996).
75. Abu-Omar, M. M., Loaiza, A. & Hontzeas, N. Reaction mechanisms of mononuclear non-heme iron oxygenases. *Chem. Rev.* **105**, 2227–2252 (2005).
76. Costas, M., Mehn, M. P., Jensen, M. P. & Que, L. Jr. Oxygen activation at mononuclear nonheme iron: Enzymes, intermediates, and models. *Chem. Rev.* **104**, 939–986 (2004).
77. Costas, M., Chen, K. & Que, L. Jr. Biomimetic nonheme iron catalysts for alkane hydroxylation. *Coord. Chem. Rev.* **200–202**, 517–544 (2000).
78. Lane, B. S. & Burgess, K. Metal-catalyzed epoxidations of alkenes with hydrogen peroxide. *Chem. Rev.* **103**, 2457–2474 (2003).
79. Tanase, S. & Bouwman, E. Selective conversion of hydrocarbons with H₂O₂ using biomimetic non-heme iron and manganese oxidation catalysts. *Adv. Inorg. Chem.* **58**, 29–75 (2006).
80. Chen, K., Costas, M. & Que, L. Jr. Spin state tuning of non-heme iron-catalyzed hydrocarbon oxidations: Participation of Fe^{III}-OOH and Fe^V=O intermediates. *Dalton Trans.* 672–679 (2002).
81. Chen, K. & Que, L. Jr. Stereospecific alkane hydroxylation by nonheme iron catalysts: Mechanistic evidence for an Fe^V=O active species. *J. Am. Chem. Soc.* **123**, 6327–6337 (2001).
82. Chen, K., Costas, M., Kim, J., Tipton, A. K. & Que, L. Jr. Olefin *cis*-dihydroxylation versus epoxidation by nonheme iron catalysts: two faces of an Fe^{III}-OOH coin. *J. Am. Chem. Soc.* **124**, 3026–3035 (2002).
This paper summarizes the mechanistic arguments in favour of the formation of the Fe(V)=O species shown in Fig. 9 that are responsible for highly stereoselective alkene oxidations by H₂O₂ that are catalysed by non-haem iron complexes.
83. White, M. C., Doyle, A. G. & Jacobsen, E. N. A synthetically useful, self-assembling MMO mimic system for catalytic alkene epoxidation with aqueous H₂O₂. *J. Am. Chem. Soc.* **123**, 7194–7195 (2001).
84. Ryu, J. Y. *et al.* High conversion of olefins to *cis*-diols by non-heme iron catalysts and H₂O₂. *Chem. Commun.* 1288–1289 (2002).
85. Mas-Ballester, R. & Que, L. Jr. Iron-catalyzed olefin epoxidation in the presence of acetic acid: Insights into the nature of the metal-based oxidant. *J. Am. Chem. Soc.* **129**, 15964–15972 (2007).
86. Mas-Ballester, R., Fujita, M., Que, L. Jr. High-valent iron-mediated *cis*-hydroxyacetoxylation of olefins. *Dalton Trans.* 1828–1830 (2008).
87. Taktak, S., Ye, W., Herrera, A. M. & Rybak-Akimova, E. V. Synthesis and catalytic properties in olefin epoxidation of novel iron(II) complexes with pyridine-containing macrocycles bearing an aminopropyl pendant arm. *Inorg. Chem.* **46**, 2929–2942 (2007).
88. Chen, M. S. & White, M. C. A predictably selective aliphatic C–H oxidation reaction for complex molecule synthesis. *Science* **318**, 783–787 (2007).
This paper shows that the non-haem iron complex Fe (18) can generate an oxidant from H₂O₂ that selectively hydroxylates particular tertiary C–H bonds in complex organic molecules.
89. Anilkumar, G., Bitterlich, B., Gelalcha, F. G., Tse, M. K. & Beller, M. An efficient biomimetic Fe-catalyzed epoxidation of olefins using hydrogen peroxide. *Chem. Commun.* 289–291 (2007).
90. Gelalcha, F. G., Bitterlich, B., Anilkumar, A., Tse, M. K. & Beller, M. Iron-catalyzed asymmetric epoxidation of aromatic alkenes using hydrogen peroxide. *Angew. Chem. Int. Edn Engl.* **46**, 7293–7296 (2007).
This paper reports the highest enantioselectivity so far for the epoxidation of an alkene by a non-haem iron catalyst.
91. Suzuki, K., Oldenburg, P. D. & Que, L. Jr. Iron-catalyzed asymmetric olefin *cis*-dihydroxylation with 97% enantiomeric excess. *Angew. Chem. Int. Edn Engl.* **47**, 1887–1889 (2008).
This paper reports the highest enantioselectivity so far for the *cis*-dihydroxylation of an alkene by a non-haem iron catalyst.
92. Lane, B. S., Vogt, M., DeRose, V. J. & Burgess, K. Manganese-catalyzed epoxidations of alkenes in bicarbonate solutions. *J. Am. Chem. Soc.* **124**, 11946–11954 (2002).
93. de Boer, J. W. *et al.* *cis*-Dihydroxylation and epoxidation of alkenes by [Mn₂O(RCO₂)₂(tmtacn)₂]: Tailoring the selectivity of a highly H₂O₂-efficient catalyst. *J. Am. Chem. Soc.* **127**, 7990–7991 (2005).
94. de Boer, J. W. *et al.* Mechanism of *cis*-dihydroxylation and epoxidation of alkenes by highly H₂O₂ efficient dinuclear manganese catalysts. *Inorg. Chem.* **46**, 6353–6372 (2007).
95. Sorokin, A. B., Kudrik, E. V. & Bouchu, D. Bio-inspired oxidation of methane in water catalyzed by N-bridged diiron phthalocyanine complex. *Chem. Commun.* 2562–2564 (2008).
This paper demonstrates the iron-catalysed oxidation of methane by H₂O₂ at relatively mild temperatures (<100 °C).
96. Collins, T. J. TAML oxidant activators: A new approach to the activation of hydrogen peroxide for environmentally significant problems. *Acc. Chem. Res.* **35**, 782–790 (2002).
This review summarizes the work of Collins's group in the use of Fe(TAML) complexes as effective catalysts for H₂O₂ activation in the treatment of waste water from paper and textile mills.
97. Chanda, A. *et al.* (TAML)Fe^V=O complex in aqueous solution: Synthesis and spectroscopic and computational characterization. *Inorg. Chem.* **47**, 3669–3678 (2008).
98. Tiago de Oliveira, F. *et al.* Chemical and spectroscopic evidence for an Fe^V-oxo complex. *Science* **315**, 835–838 (2007).
99. Balasubramanian, R. & Rosenzweig, A. Structural and mechanistic insights into methane oxidation by particulate methane monooxygenase. *Acc. Chem. Res.* **40**, 573–580 (2007).

Acknowledgements Work on biologically inspired oxidation catalysis carried out in the laboratory of L.Q. is supported by the US Department of Energy, and support for copper oxidation chemistry in the laboratory of W.B.T. is provided by the National Institutes of Health.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to the authors (larryque@umn.edu; wtolman@umn.edu).

Towards uranium catalysts

Alexander R. Fox¹, Suzanne C. Bart², Karsten Meyer² & Christopher C. Cummins¹

The forefront of research into the complexes of uranium reveals chemical transformations that challenge and expand our view of this unique element. Certain ligands form multiple bonds to uranium, and small, inert molecules such as nitrogen and carbon dioxide become reactive when in complex with the metal. Such complexes provide clues to the catalytic future of uranium, in which the applications of the element extend far beyond the nuclear industry. Most excitingly, the ability of uranium to use its outermost *f* electrons for binding ligands might enable the element to catalyse reactions that are impossible with conventional, transition-metal catalysts.

The non-aqueous chemistry of uranium has been an active area of exploration in recent decades^{1,2}. The results of these investigations paint a picture of the heaviest naturally abundant element that reflects its unique position both in the periodic table and in modern life. No other element has such a Jekyll and Hyde reputation, associated as uranium is with the potential for either a carbon-free energy supply or, alternatively, mass destruction. Beyond applications to nuclear power, future beneficial roles for uranium will be created in part by the quest of researchers to understand the properties and potential applications of uranium-containing molecules and materials. As humanity intensifies its search for carbon-neutral energy sources, the demand for enriched uranium as a fissile nuclear fuel can be expected to rise. With this comes a grand opportunity for researchers to develop new and beneficial uses for the high-purity depleted uranium produced as a by-product of nuclear isotope enrichment programmes.

The early actinide elements, uranium included, have features that distinguish them from the transition-metal elements and lanthanide elements (see page 314). For example, the comparatively large ionic radius of uranium and the availability of *f* orbitals to engage in bonding interactions generate coordination numbers and modes not commonly observed for the lighter lanthanide or transition elements. Additionally, the range of stable oxidation states that uranium can access, including high-valent uranium(V) and uranium(VI), further distinguishes it from the more often trivalent lanthanides. The result is a palette of chemical reactivity that is special to uranium and complementary to that of the *d*-block transition elements. As a result, complexes featuring uranium have begun to find applications in chemical catalysis³. Examples of catalytic processes mediated by uranium complexes include hydrogenation of alkenes^{4–8}, oligomerization, dimerization, hydrosilation, and hydroamination of terminal alkynes^{9–11}. Many of these catalytic processes were first demonstrated by stoichiometric reactions, whose further study allowed catalysis to be realized.

This article is concerned with highlighting the forefront of coordination chemistry, small-molecule activation and atom/group-transfer reactions with well-defined molecular uranium complexes, with an eye towards catalysis with uranium. From classical goals such as carbon–hydrogen bond activation to recent and unprecedented reactivity involving simple chemical feedstock such as dinitrogen and carbon monoxide, delineating the state of the art of synthetic uranium chemistry should provide form to its future in catalysis, drawn from an expanding array of known reactions promoted by the element with atomic number 92.

Dinitrogen binding and reduction

The synthesis of ammonia from the elements N₂ and H₂, as practised in the Haber–Bosch process, requires harsh conditions of temperature and pressure. These requirements can be traced to the great strength of the N₂ triple bond, the activation of which by *d*-block elements has been closely studied with the objective being a milder process for nitrogen

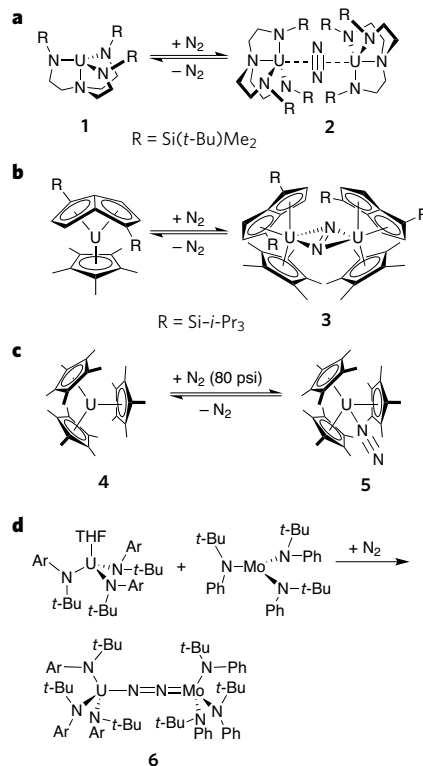


Figure 1 | Uranium dinitrogen complexes. Complexes of uranium have been shown to bind to N₂ in several ways: in a bridging, side-on manner (**a**, **b**); in a terminal, end-on manner (**c**); and in a bridging, end-on manner (**d**). Although the degree of activation of the coordinated N₂ ligand varies, only in **6** is the binding of N₂ irreversible, owing to a highly reducing molybdenum centre working together with uranium. Ar, 3,5-dimethylphenyl; *i*-Pr, isopropyl; *t*-Bu, *tert*-butyl; Me, methyl; THF, tetrahydrofuran.

¹Department of Chemistry, Room 6-435, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139-2307, USA. ²Department of Chemistry and Pharmacy, Inorganic Chemistry, Friedrich Alexander University Erlangen-Nuremberg, Egerlandstrasse 1, Erlangen, Bavaria 91058, Germany.

fixation. Early catalysts used in the Haber–Bosch process contained uranium, demonstrating the reducing power of this element with respect to small, inert molecules such as dinitrogen¹².

One of the first actinide dinitrogen complexes was synthesized by exposure of the trivalent complex $[\text{U}(\text{N}_3\text{N})]$ (**1**; $\text{N}_3\text{N} = \text{N}(\text{CH}_2\text{CH}_2\text{NSi}(t\text{-Bu})\text{Me}_2)_3$; $t\text{-Bu} = -\text{C}(\text{CH}_3)_3$; $\text{Me} = -\text{CH}_3$) to N_2 , producing the side-on bound dinitrogen complex $[(\text{N}_3\text{N})\text{U}]_2(\mu, \eta^2, \eta^2\text{-N}_2)$ (**2**), in which the N_2 ligand is sandwiched between two trigonal monopyramidal uranium centres¹³ (Fig. 1a). The N–N bond length of $1.109 \pm 0.007 \text{ \AA}$ is essentially that of free dinitrogen (1.0975 \AA), indicating little activation in this example of a side-on coordinated dinitrogen ligand. Reduction of dinitrogen did occur in the side-on bound uranium(IV) dinitrogen pentalene complex $[(\text{Cp}^*)(\eta^5\text{-}1,4\text{-}(\text{Si-}i\text{-Pr}_3)_2\text{C}_8\text{H}_4)\text{U}]_2(\mu, \eta^2, \eta^2\text{-N}_2)$ (**3**; $\text{Cp}^* = \eta^5\text{-C}_5\text{Me}_5$; $i\text{-Pr} = -\text{CH}(\text{CH}_3)_2$). Here, the N–N bond length of $1.232 \pm 0.010 \text{ \AA}$ is consistent with an N–N double bond resulting from partial reduction of the bound dinitrogen by the uranium centre¹⁴ (Fig. 1b).

Several examples of end-on bound dinitrogen complexes have also been observed. The metallocene complex $[(\text{Cp}^*)_3\text{U}]$ (**4**) reversibly binds dinitrogen in an end-on manner. The resultant complex, $[(\text{Cp}^*)_3\text{U}(\eta^1\text{-N}_2)]$ (**5**), has a non-activated N_2 ligand as supported by both X-ray crystallography and infrared spectroscopy (Fig. 1c). Both the N–N distance of $1.120 \pm 0.014 \text{ \AA}$ and the band for the N–N stretch observed at $2,207 \text{ cm}^{-1}$ ($^{15}\text{N}_2$ analogue = $2,134 \text{ cm}^{-1}$) indicate that there is no reduction by the uranium centre. A heterodinuclear U–Mo compound with a bridging μ, η^1, η^1 -dinitrogen ligand, $[(\text{Ar}[t\text{-Bu}]\text{N})_3\text{U}(\mu, \eta^1, \eta^1\text{-N}_2)\text{Mo}(\text{N}[t\text{-Bu}]\text{Ph})_3]$ (**6**; $\text{Ar} = 3,5\text{-Me}_2\text{C}_6\text{H}_3$; $\text{Ph} = \text{phenyl}$), was synthesized by exposing an equimolar mixture of $[(\text{Ar}[t\text{-Bu}]\text{N})_3\text{U}(\text{THF})]$ and $[\text{Mo}(\text{N}[t\text{-Bu}]\text{Ph})_3]$ in toluene to N_2 at 1 atmosphere (ref. 15) (Fig. 1d). The highly reactive trivalent precursor, $[(\text{Ar}[t\text{-Bu}]\text{N})_3\text{U}(\text{THF})]$, is thought to combine with an intermediate dinitrogen complex $[(\text{Ph}[t\text{-Bu}]\text{N})_3\text{Mo}(\eta^1\text{-N}_2)]$ formed in the course of the reaction. Synthesis of **6** with $^{15}\text{N}_2$ revealed a band in the infrared spectrum at $1,547 \text{ cm}^{-1}$ that was assigned to the $^{15}\text{N}\text{--}^{15}\text{N}$ stretching mode. The structure of **6** revealed activation of the N_2 fragment, because the N–N distance was decreased from that of free dinitrogen to $1.232 \pm 0.011 \text{ \AA}$. The Mo–N distance of $1.773 \pm 0.008 \text{ \AA}$ is consistent with a multiple-bond interaction¹⁶, and the U–N distance of $2.220 \pm 0.009 \text{ \AA}$ suggests some degree of multiple bonding between the N_2 fragment and the uranium centre.

Remarkably, even cleavage of dinitrogen was achieved by treating the uranium salt $[(\text{Et}_6\text{-calix}[4]\text{tetrapyrrole})\text{U}(\text{dme})][\text{K}(\text{dme})]$ ($\text{dme} = 1,2\text{-dimethoxyethane}$; $\text{Et} = -\text{CH}_2\text{CH}_3$) with one equivalent of potassium naphthalenide in dme in the presence of dinitrogen gas¹⁷. The result is a mixed-valent μ -nitrido U(V/IV) complex. This result shows the ability of a highly reducing uranium centre, when properly ligated and used in combination with an external reducing agent, to reduce the strong dinitrogen triple bond completely.

The side-on coordination mode of dinitrogen typically results in a greater N–N bond activation by donation of electron density from metal-based π -symmetric orbitals into the N_2 -based π^* -orbital^{18–20}. Consequently, functionalization of η^2 -dinitrogen ligands by methods such as hydrogenation may be achieved more readily. The labile nature of the dinitrogen ligands in **2** and **3** may present a challenge to the synthetic elaboration of the uranium-bound N_2 ligands, although simply conducting reactions under a higher pressure of N_2 may circumvent such issues. End-on coordination of dinitrogen tends to result in weaker metal– N_2 interactions, owing to the poor σ -donor and π -acceptor characteristics of N_2 . Coupling $\eta^1\text{-N}_2$ ligands with external reducing equivalents can promote further N_2 activation and functionalization, and the potential of this strategy has been demonstrated with the catalytic production of ammonia by N_2 activation by using a molybdenum-based catalyst²¹.

Binding of carbon monoxide and carbonylation chemistry

Since the development of the Fischer–Tropsch process in the 1930s, the activation of carbon monoxide (CO) has been explored for the production of synthetic petroleum with a variety of transition-metal complexes. Despite this fact, little is known about the reactivity and activation of this small C_1 molecule by actinide elements. Because uranium

can form and stabilize charge-separated species, it has potential for the activation and functionalization of CO.

The earliest reported uranium complex of CO resulted from the exposure of $[(\eta^5\text{-Me}_3\text{SiC}_5\text{H}_4)_3\text{U}]$ to CO gas²². Formation of the carbonyl complex was confirmed by the CO absorption band in the infrared spectrum centred at $1,976 \text{ cm}^{-1}$, present at $1,935 \text{ cm}^{-1}$ when ^{13}CO gas is used and at $1,932 \text{ cm}^{-1}$ when C^{18}O is used. Application of vacuum to the compound in solution regenerates the starting material, precluding isolation of the product. By analogy with the corresponding isonitrile compound, $[(\eta^5\text{-Me}_3\text{SiC}_5\text{H}_4)_3\text{U}(\text{CNet})]$, the CO complex $[(\eta^5\text{-Me}_3\text{SiC}_5\text{H}_4)_3\text{U}(\text{CO})]$ was thought to contain a C-bound carbonyl ligand and a linear U–C–O unit.

After this initial report, the closely related uranium carbonyl complex $[(\eta^5\text{-Me}_4\text{C}_5\text{H}_5)_3\text{U}(\eta^1\text{-CO})]$ was isolated and characterized spectroscopically and crystallographically²³. Its infrared spectrum features an intense C–O stretch at $1,900 \text{ cm}^{-1}$, indicating the presence of significant U–CO π -backbonding. X-ray crystallography showed a linear U–C–O angle ($175.2 \pm 0.6^\circ$) and a short U–C bond ($2.383 \pm 0.006 \text{ \AA}$), also indicative of a strong bonding interaction with the uranium centre. The C–O distance of $1.142 \pm 0.007 \text{ \AA}$ is similar to that in $[(\text{Cp}^*)_2\text{Ti}(\text{CO})_2]$ of 1.149 \AA (ref. 24).

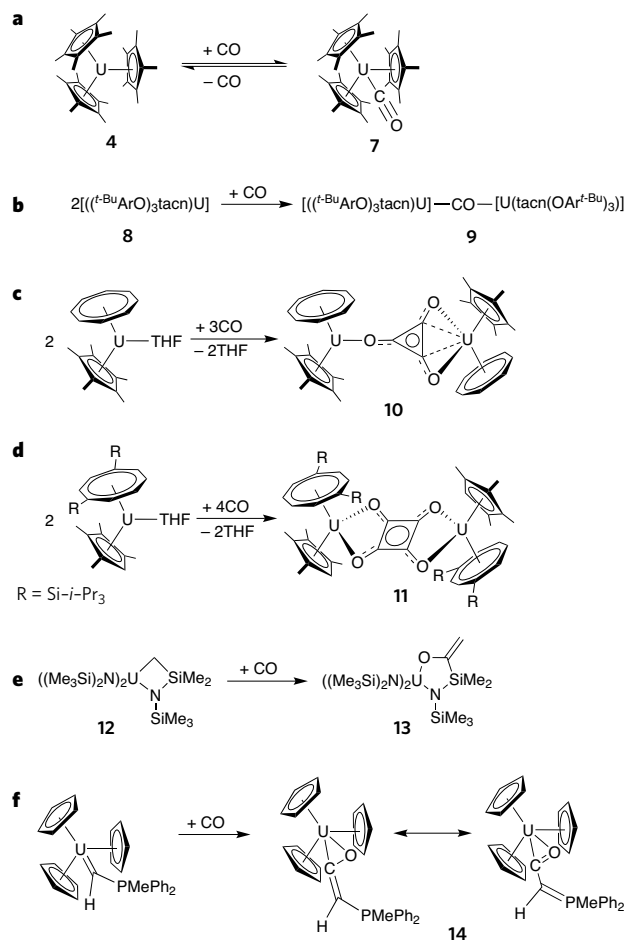


Figure 2 | Binding and reductive coupling of carbon monoxide. **a**, Early examples of CO complexation by uranium were typified by weak, reversible binding of CO in a terminal, end-on manner. **b**, Recent work has shown that CO can bind in ways previously unseen within the coordination sphere of uranium, as is the case for the isocarbonyl ligand that bridges two uranium centres in **9** ($(t\text{-BuArO})_3\text{tacn}^{3-}$ = trianion of 1,4,7-tris(3,5-di-*t*-butyl-2-hydroxybenzyl)-1,4,7-triazacyclononane). **c**, **d**, Carefully constructing the coordination environment of uranium has allowed the reductive coupling of CO molecules by uranium(III) ions to form the deltate anion, $(\text{C}_3\text{O}_3)^{2-}$ (**c**) and the squarate anion, $(\text{C}_4\text{O}_4)^{2-}$ (**d**). **e**, **f**, CO has also been shown to insert into reactive ligand functionalities such the U–C single (**e**) and double (**f**) bonds. Ph, phenyl.

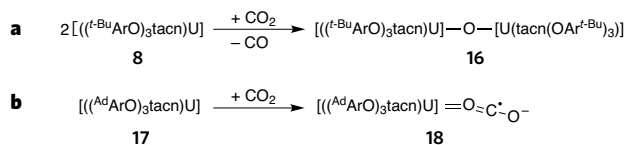


Figure 3 | Coordination and activation of carbon dioxide. **a**, Abundant and typically inert, CO₂ can be reduced by two electrons by the reactive uranium(III) complex **8**, releasing CO and producing a bridging oxide complex (**16**). This reaction probably takes place through the initial formation of a CO₂ adduct of **8**, following which a second equivalent of **8** enters to complete the reduction process and form **16**. **b**, For **17**, in which the *t*-butyl groups of the ancillary ligand set have been replaced by 1-adamantyl (Ad) groups, isolation of a CO₂ adduct of uranium occurs rather than cleavage of the C–O bond. In **18**, the CO₂ ligand binds in an unprecedented terminal, end-on, manner and has been reduced by one electron. The increased steric bulk of the ligand set, in comparison with that of **8**, prohibits the formation of a bimolecular complex, thus preventing CO elimination.

Adding another methyl group to the cyclopentadienyl ligand produced the similar uranium carbonyl species, [(Cp^{*})₃U(η¹-CO)] (ref. 25) (**7**, Fig. 2a), with an infrared spectroscopic absorption at 1,922 cm⁻¹, which shifts to 1,877 cm⁻¹ with ¹³CO. The linearly bound CO ligand (U–C–O angle of 180°) is crystallographically disordered, with the CO lying on an axis of symmetry. The resultant C–O bond distance of 1.13±0.01 Å indicates little activation²⁶. Addition of CO to [(^t-BuArO)₃tacn]U (**8**; (^t-BuArO)₃tacn³⁻ = trianion of 1,4,7-tris(3,5-di-*t*-butyl-2-hydroxybenzyl)-1,4,7-triazacyclononane) afforded the first example of a CO-bridged dinuclear uranium species, {[(^t-BuArO)₃tacn]U₂(μ,η¹,η¹-CO)} (ref. 27) (**9**, Fig. 2b). The reaction is thought to occur by nucleophilic attack of a charge-separated [(^t-BuArO)₃tacn]U(IV)–CO⁻ fragment on the coordinatively unsaturated uranium(III) centre in **8**. X-ray diffraction revealed a bridging end-on coordination mode of the isocarbonyl ligand, with a short U–C bond and a longer U–O interaction. Crystallographic disorder of the carbonyl ligand, however, prevents the determination of reliable distances of the bridging CO ligand and consequently precludes assessing the degree of activation by this metric.

The previously mentioned complexes have demonstrated the first important step in a metal-mediated catalytic process, the ability of the uranium ion to coordinate and activate the substrate (CO). Subsequent formation of new bonds has been realized by addition of CO to a family of metallocene-inspired uranium complexes. Exposure of the mixed metallocene complex [(η⁸-COT)(Cp^{*})U(THF)] (COT = 1,3,5,7-cyclooctatetraene) to CO produces a dinuclear deltate (C₃O₃²⁻) uranium(IV) complex, {[(η⁸-COT)(Cp^{*})U]₂(μ,η¹,η²-C₃O₃)} (ref. 28) (**10**, Fig. 2c). The resultant planar deltate core in **10** arises from uranium-mediated C–C bond formation between three CO molecules. The C–O bond distances are intermediate between those of single and double bonds. There are two short C–C distances and one long distance, the latter featuring an interaction with one uranium centre. Modifying the steric demand of the supporting ligands provides access to the corresponding dinuclear squarate (C₄O₄²⁻) derivative (ref. 29) (**11**, Fig. 2d).

Carbon–carbon bond formation has also been demonstrated by the insertion of CO into uranium–carbon bonds. Carbon monoxide inserts into the U–C alkyl bond of [(η⁵-MeC₅H₄)₃U(*t*-Bu)], creating [(η⁵-MeC₅H₄)₃U(C(O)-*t*-Bu)], which contains an η²-bound acyl group³⁰. The related complexes, [(Cp)₃U(R)] (R = Me, Et, *i*-Pr, *n*-Bu, *t*-Bu; *n*-Bu = –CH₂CH₂CH₂CH₃), store CO in the form of the corresponding η²-acyl complexes, [(Cp)₃U(C(O)R)] (ref. 31). Clean extrusion of CO from [(Cp)₃U(C(O)R)] is achieved on being heated to 60 °C. Although clean insertion chemistry with CO was not observed for the triamide-supported uranium methyl complex, [(Me₃Si)₂N]₂U(Me)] (ref. 32), the related metallocyclic species, [(Me₃Si)₂N]₂U(CH₂Si(Me)₂NSiMe₃) (**12**), reacts with CO to generate a new metallocyclic derivative (**13**), which features an exocyclic vinyl group derived from CO insertion and methylene migration³³ (Fig. 2e). The polarized U–C bond and inherent ring strain in **12**, coupled with the strength of the U–O bond in **13**,

provide a thermodynamic driving force for this chemistry. Insertion of CO into the U–C multiple bond of [(Cp)₃U(=CHPMePh₂)] also generates the tetrahedral β-ketoylide derivative [(Cp)₃U(η²-COCHPMePh₂)] (**14**), featuring an η²-CO fragment³⁴. In **14**, the C–O bond length of 1.27±0.03 Å, the P–C distance of 1.77±0.02 Å and the C–C distance of 1.37±0.03 Å suggest a delocalized structure modelled by two resonance structures (Fig. 2f). Carbonylation chemistry has also been described with the insertion of CO into uranium–nitrogen bonds³¹.

Clearly, CO can interact with uranium complexes in a variety of modes. The ability of uranium not only to activate the coordinated CO fragment but also to functionalize it indicates that uranium may be a viable metal for catalytic CO transformations, provided that the carbonylated organic fragments can be removed from the metal centre.

Complexes and reduction of carbon dioxide

Carbon dioxide (CO₂) is a favourable C₁ source because of its low cost and availability, but its thermodynamically stable double bonds must be reduced for it to serve as a practical building block. Transition-metal catalysts have been used for CO₂ activation^{35,36}, but using uranium instead may offer previously unavailable modes for fixation of this key small molecule.

Derivatization of CO₂ has been accomplished with the formation of new carbon–element bonds. The uranium(IV) carbamate species [U(O₂CNEt₂)₄] was generated by means of salt elimination by the addition of a premixed solution of diethylamine and CO₂ containing diethylammonium diethylcarbamate to a toluene solution of UCl₄ (ref. 37). Insertion of CO₂ into the uranium–sulphur bond of the uranium(IV) dithiolate complex, [(Cp^{*})₂U(SR)₂] (R = *t*-Bu), forms new carbon–sulphur bonds in the complex [(Cp^{*})₂U(O₂CSR)₂] (ref. 38). Thermolysis of [(Cp^{*})₂U(O₂CSR)₂] elicits CO₂ extrusion and the formation of [(Cp^{*})₂U(SR)₂]. Tobin Marks demonstrated the insertion of CO₂ into the uranium–carbon bonds of the actinide dimethyl complex, [(Cp^{*})₂U(Me)₂] (**15**) (ref. 39). The bis(acetate) complex, [(Cp^{*})₂U(OAc)₂], was formed by subjecting **15** to CO₂ at 1 atmosphere³⁹. Infrared spectroscopy and determination of the molecular mass in solution suggested a monomeric complex containing bidentate acetate ligands. Careful addition of only one equivalent of CO₂ to **15** formed the monomeric methyl acetate complex, [(Cp^{*})₂U(Me)(OAc)], which has one bidentate acetate ligand. These examples of CO₂ activation are similar to what is observed for corresponding *d*-block element systems; on moving to uranium(III) interactions with CO₂, however, a significant departure is observed.

Addition of CO₂ to trivalent **8** results in a two-electron reduction of CO₂ to release CO and produce an oxide-bridged U(IV/IV) species, {[(^t-BuArO)₃tacn]U₂(μ-O)} (ref. 27) (**16**, Fig. 3a). Replacing the *ortho t*-butyl substituents on the coordinating aryl oxides with adamantyl groups forms [(^{Ad}ArO)₃tacn]U (**17**; Ad = 1-adamantyl), which, on exposure to CO₂, generates a previously unseen end-on CO₂–uranium complex [(^{Ad}ArO)₃tacn]U(η¹-OCO) (ref. 40) (**18**, Fig. 3b). In this unusual complex the CO₂ ligand is bound almost linearly, with U–O–C and O–C–O angles of 171.1±0.2° and 178.0±0.3°, respectively. The vibrational band for the CO₂ ligand in **18** appears at 2,188 cm⁻¹ (2,128 cm⁻¹ for ¹³CO₂), indicating that this fragment is reduced in comparison with free CO₂ (2,349 cm⁻¹) (ref. 41). Reduction is further supported by crystallographic evidence showing that the coordinated and terminal O–C distances differ. The O–C bond length of 1.122±0.004 Å for the unit directly coordinated to the uranium centre is in the range expected for a double bond. However, the terminal O–C bond distance of 1.277±0.004 Å is much longer, indicating that activation has transpired. The crystallographic evidence, supported by spectroscopic studies, confirms the formulation of **18** as a charge-separated uranium(IV) species, U(IV)–L⁻ (L = ligand), with a resonance-stabilized radical within the U–OCO unit⁴⁰. This exemplified ability of uranium to form stable charge-separated species may be an important mediator for the stabilization of reactive intermediates in catalytic transformations and is a distinguishing factor from *d*-block element chemistry.

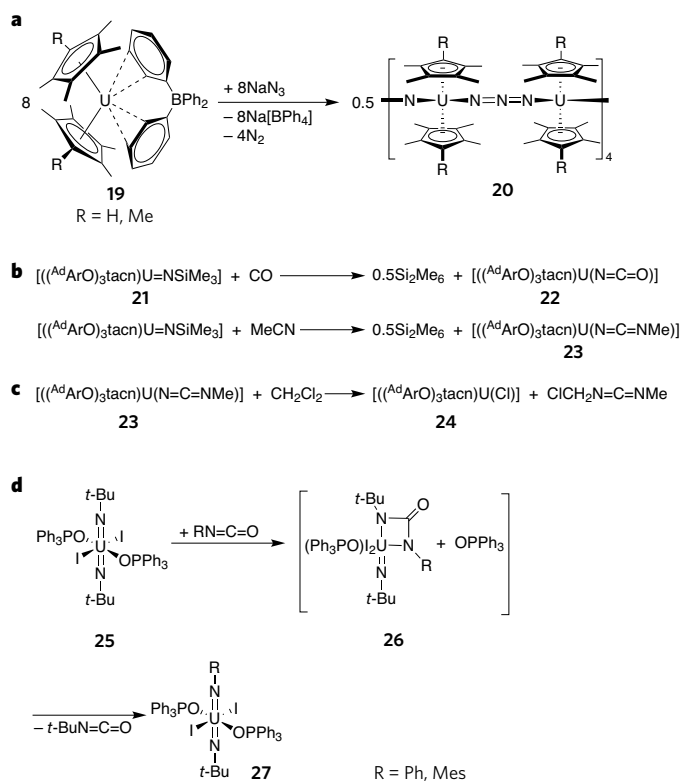


Figure 4 | Azido, imido and nitrido complexes of uranium. The potential application of binary uranium nitride as a nuclear fuel has spurred interest in soluble molecular precursors that contain the extremely rare uranium-nitride functionality. **a**, One such example is **20**, in which eight uranium centres, four azide bridges and four nitride bridges form nanometre-sized rings. **b**, Imido complexes of uranium are of interest because of the unique reactivity patterns derived from the U–N multiple bond. The uranium(v)–imido complex **21** engages in multiple-bond metathesis with CO or acetonitrile (MeCN) to form the corresponding uranium(iv) derivatives **22** and **23** together with Si_2Me_6 . **c**, The chemistry of **23** has been extended to include the synthesis of organic carbodiimide molecules by means of a closed stoichiometric cycle. **d**, The multiple-bond metathesis reactivity of uranium–imido complexes has been furthered by the recently reported *trans*-bis(imido)uranium(vi) complex (**25**), which reacts with aryl isocyanates to exchange imido ligands through a proposed metalocyclic intermediate (**26**). Mes, 2,4,6-trimethylphenyl.

Activation of azides and amines

Reactions of azides (as RN_3 or N_3^- salts) with metal complexes have been explored extensively because they represent a frequent entry point to the synthesis of metal imido or nitrido complexes. Typical metal azido species can be converted to nitrido fragments through photolysis or thermolysis to facilitate loss of dinitrogen. Uranium nitride complexes are of particular interest because the expected polarity within the U–N moiety should impart enhanced reactivity. This attribute may facilitate the synthesis of N-functionalized small molecules within the coordination sphere of uranium by means of N-atom transfer reactions or aziridination⁴². Uranium azide and nitride species are also significant because of their proposed use in nuclear fuel cycles. Binary uranium nitrides have advantages over other nuclear fuels because they are dense, uranium-rich materials with high thermal conductivities and can function at high temperatures⁴³.

Complexes featuring the uranium nitride functionality have been synthesized recently in the form of unusual uranium–nitrogen rings with the core formula $(\text{UNUN}_3)_4$. Treatment of NaN_3 with the metallocene derivative $[(\eta^5\text{-C}_5\text{Me}_4\text{R})_2\text{U}(\mu\text{-Ph})_2\text{BPh}_2]$ (**19**, R = Me, H) produces nanometre-sized rings with the formula $[(\eta^5\text{-C}_5\text{Me}_4\text{R})_2\text{U}(\mu\text{-N})(\eta^5\text{-C}_5\text{Me}_4\text{R})_2\text{U}(\mu\text{-N}_3)]_4$ (**20**) containing alternating nitrido and azido linkages⁴⁴ (Fig. 4a). The nitride bridges have short U–N distances within the double-bond length range (from 2.047 ± 0.006 Å to 2.090 ± 0.008 Å). Infrared spectroscopy

identifies the azide ligands with an intense absorption at $2,100\text{ cm}^{-1}$. The formation of **20** occurs by reduction of four azido units by eight uranium(III) centres to afford four nitrido ligands. Mixed uranium azide/nitride cluster formation has also been demonstrated by the reaction of $\text{U}_3(\text{THF})_4$ with the unstable uranium hepta-azido salt, $\text{Cs}_3[\text{U}(\text{N}_3)_7]$, in acetonitrile, which led to the isolation of a one-dimensional polymer of unit formula $\text{Cs}[\text{U}_4(\mu_4\text{-N})(\mu\text{-N}_3)_8(\text{CH}_3\text{CN})_7\text{I}_6]$ (refs 45, 46). At the core of each monomeric unit is a distorted tetrahedron composed of four uranium(IV) ions, one at each vertex, a nitride ligand that occupies a bridging position at the approximate centre of the tetrahedron, and eight azide ligands that bridge the uranium vertices in an end-on manner. Remaining coordination sites at uranium are occupied by iodide and acetonitrile, whereas caesium cations enforce the polymeric structure by bridging iodide ligands of adjacent tetrahedral clusters. These azide derivatives can be compared to the classic monomeric metallocene-based uranium bis(azide), $[(\text{Cp}')_2\text{U}(\text{N}_3)_2]$ ($\text{Cp}' = \eta^5\text{-1,2,4-(t-Bu)}_3\text{C}_5\text{H}_2$) (ref. 47).

Azide activation is also observed on addition of organic azides to low-valent uranium complexes, which additionally generates high-valent monomeric uranium imido complexes through N_2 extrusion. For instance, treating $[(\eta^5\text{-MeC}_5\text{H}_4)_3\text{U}]$ with phenyl azide resulted in N_2 loss to produce the uranium(v) imido species $[(\eta^5\text{-MeC}_5\text{H}_4)_3\text{U}(\text{NPh})]$, where the imido ligand engages in a formal triple bond to uranium⁴⁸. A similar reaction produced the first uranium(vi) bis(imido) complex, $[(\text{Cp}')_2\text{U}(\text{NPh})_2]$ (ref. 49). Similar bis(imido) derivatives were synthesized by the reductive cleavage of 1,2-disubstituted hydrazines⁵⁰ and the direct reduction of diazene and azide molecules⁵¹. Non-cyclopentadienyl ancillary ligands can support uranium(v) imido complexes as well. Use of the chelating triamidoamine ligand system, $[\text{N}_3\text{N}]$, allowed the synthesis of the uranium(v) trimethylsilylimido derivative, $[(\text{N}_3\text{N})\text{U}(\text{NSiMe}_3)]$ (ref. 52), and a similar non-chelating ligand framework generated the analogous derivative, $[(\text{Me}_3\text{Si})_2\text{N}_3\text{U}(\text{NSiMe}_3)]$, which was characterized by X-ray crystallography⁵³. This molecule has a very short U–N(imido) bond distance of 1.91 ± 0.02 Å, indicative of a formal triple bond.

Addition of azidotrimethylsilane to **8** or **17** in benzene affords the uranium(v) imido compounds, $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\text{NSiMe}_3)]$ and $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\text{NSiMe}_3)]$ (**21**)^{54,55}. Crystallographic analysis of $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\text{NSiMe}_3)]$ revealed a U–N(imido) distance of 1.989 ± 0.005 Å and a U–N–Si angle of $173.7 \pm 0.3^\circ$, again consistent with a U–N triple bond. The more sterically hindered **21** has the longest ever reported U–N(imido) bond distance, 2.1219 ± 0.0018 Å. The U–N–Si angle of $162.55 \pm 0.12^\circ$ deviates significantly from linearity, a fact that is most probably due to the steric pressure imparted by the 1-adamantyl groups, pressure that prevents the imido nitrogen *p* orbitals from engaging in optimal metal–ligand π -bonding. This causes increased nucleophilicity at the imido nitrogen and imparts reactivity towards π -acids. Thus, treatment of **21** with CO or MeCN produces the respective uranium(iv) isocyanato and carbodiimido complexes, $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\eta^1\text{-NCO})]$ (**22**) and $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\eta^1\text{-NCNMe})]$ (**23**), by means of multiple bond metathesis and release of Si_2Me_6 (ref. 55) (Fig. 4b). Interestingly, the carbodiimide can be released from the uranium centre in **23** by addition of methylene chloride, generating the uranium(iv) chloride, $[(^{\text{Ad}}\text{ArO})_3\text{tacn})\text{U}(\text{Cl})]$ (**24**), along with $\text{MeN}=\text{C}=\text{NCH}_2\text{Cl}$ (Fig. 4c). Reduction of **24** regenerates **17** for reoxidation by azidotrimethylsilane.

James Boncella and co-workers recently reported the formation of previously unknown *trans*-bis(imido)uranium(vi) complexes featuring the core $[\text{U}(\text{NR})_2]^{2+}$ fragment, analogues of the uranyl ion, $[\text{UO}_2]^{2+}$. Addition of *t*-BuNH₂ and I₂ to elemental uranium turnings in THF created the octahedral *trans*-diimide complex, $[(t\text{-BuN})_2\text{UI}_2(\text{THF})_2]$ ($\text{U}-\text{N}_{\text{avg}} = 1.84$ Å) with two THF and two iodide ligands occupying the equatorial plane in a *cis* manner⁵⁶. Similar *trans*-diarylimide complexes are obtained by the addition of RNH_2 (R = Ph, 3,5-(CF₃)₂C₆H₃, 2,6-(*i*-Pr)₂C₆H₃) and I₂ to $[\text{U}_3(\text{THF})_4]$ in THF⁵⁷. The reactivity of $[(t\text{-BuN})_2\text{UI}_2(\text{THF})_2]$ includes simple substitution reactions such as exchanging the coordinated THF for a more potent donor ligand such as Ph_3PO , or iodide replacement with trifluoromethanesulphonate

by reaction with AgOTf (Tf = CF₃SO₂) (ref. 57). Reactivity of the imido functionality was demonstrated with the synthesis of a mixed *trans*-imido(oxo) complex, [(*t*-BuN)(O)U(THF)₂], by reaction of [(*t*-BuN)₂U(THF)₂] with [B(C₆F₅)₃·H₂O] (ref. 58). Furthermore, the phosphine oxide derivative, [(*t*-BuN)₂U(OPPh₃)₂] (**25**), reacts with one or two equivalents of aryl isocyanates, RNCO (R = Ph or 2,4,6-trimethylphenyl (Mes)), to provide the mixed *trans*-diimide complex [(*t*-BuN)(RN)U(OPPh₃)₂] (**27**) and the symmetric *trans*-diimide complex [(RN)₂U(OPPh₃)₂], respectively⁵⁹. Calculations suggest a reaction mechanism involving [2 + 2] cycloaddition of the aryl isocyanate C–N double bond across the uranium imido linkage, forming an *N,N'*-bound metallocyclic intermediate (**26**) that then liberates the **27** and *t*-BuNCO through retro [2 + 2] cycloaddition (Fig. 4c).

Hydrocarbon activation

Activation of simple alkanes to form more complicated, desirable products continues to be a challenge in catalysis. The strong C–C and C–H bonds often must be activated and functionalized for elaboration into industrially relevant molecules. Previous work has shown that actinides such as thorium are capable of such reactivity through a σ -bond metathesis route^{60,61}, and more recently activation with uranium has been observed.

Intramolecular C–H bond activation was demonstrated by addition of diphenylditelluride to [(Cp*)₂U(Me)₂] (**15**), the reaction resulting in loss of methane and PhTeMe to produce [(Cp*)₂U(η^2 -(Te,C)-TeC₆H₄)] (ref. 62) (**28**, Fig. 5a). Activation of one phenyl C–H bond produces a nearly planar four-membered U–Te–C–C ring. The U–C distance is 2.352 ± 0.004 Å and is within the range expected for uranium–aryl single bonds. Similarly, addition of pyridine-*N*-oxide to **15** or the dibenzyl derivative, [(Cp*)₂U(CH₂Ph)₂] (**29**, ref. 63), produced one equivalent of

either methane or toluene, respectively, and a uranium(IV) centre coordinated in an η^2 -manner by an activated pyridine-*N*-oxide^{64–66} (Fig. 5b). One alkyl group is retained, producing the cyclometallated complex [(Cp*)₂U(R)(η^2 -(O,C)-ONC₅H₄)] (**30**, R = Me; **31**, R = CH₂Ph).

Similar C–H activation to form substituted indazoles has been shown with the sterically pressured uranium diazoalkane complex [(^{Ad}ArO)₃tacn)U(N₂CPh₂)] (ref. 67). Heating this complex, which features an η^1 -bound diphenyldiazomethane ligand, results in cleavage of the *ortho*-C–H bond of one of the phenyl substituents followed by nitrogen insertion into the new U–C bond, producing the phenyl-substituted indazole uranium complex, [(^{Ad}ArO)₃tacn)U(η^2 -3-Ph-Ind)] (Ind = C₇H₅N₂, indazole).

Intramolecular C–H bond activation by uranium ions has been coupled with isotopic labelling in the triamido-based uranium(IV) system, [(Me₃Si)₂N₃U(H)] (**32**) (refs 68, 69). Under an atmosphere of D₂, **32** undergoes complete H–D exchange of the metal-bound hydride ligand and all 54 of the hydrogen atoms of the ancillary ligand set. The reaction may be completely reversed by subjecting the perdeutero derivative of **32** to an atmosphere of H₂. The hydride ligand in **32** can exchange with H₂ or D₂ by simple σ -bond metathesis, but H–D exchange at the hexamethyldisilazide ligands must take place with C–H bond activation. Pyrolysis of **32** provides the C–H activated metallocycle, **12**, through loss of H₂. Exposure of **12** to D₂ leads to the perdeutero derivative of **32**, thus implicating **12** as a key intermediate in the H–D exchange process. A series of equilibrium steps involving elimination of H₂ (or D₂) from **32**, formation of **12** and subsequent reaction with D₂ (or H₂) accounts for the observed transformation and its reversibility (Fig. 5c).

An example of C–H bond coordination to uranium was demonstrated by addition of cyclic alkanes to **8**, providing the first crystallographically characterized example of a C–H bond interacting with a uranium centre⁷⁰. The uranium–methylcyclohexane complex [(^t-BuArO)₃tacn)U(C₇H₁₄)] (C₇H₁₄ = methylcyclohexane) has a uranium–carbon distance of 3.864 Å and short contacts between the peripheral *t*-butyl groups and the axial alkane ligand (from 2.12 Å to 2.71 Å). The uranium–carbon distances are less than the sum of the van der Waals radii for a U–CH₂ or U–CH₃ contact (3.9 Å) and suggest a bonding interaction between the alkane and the U centre. X-ray diffraction analysis at 5 K allowed the refinement of all hydrogen atoms and established an η^2 -C,H coordination of the alkane ligand. This snapshot of an alkane interacting with a uranium centre may provide insight into the type of molecular architecture needed for catalytic C–H activation with uranium ions.

Reactions consuming or producing H₂

The binary uranium(III) hydride, UH₃, reduces select substrates, such as PhBr, PhNO₂ and Ph₃AsO, in moderate to good yields; however, the polymeric nature and consequent low solubility of UH₃ limit its utility for most synthetic applications⁷¹. Complexes featuring reactive U–H bonds offer advantages over UH₃ such as increased solubility, and have therefore been applied more extensively. The synthetic utility of hydride complexes derived from the [bis(Cp)uranium] framework has been clearly demonstrated⁹. These compounds perform stoichiometric and catalytic hydrogenation of unsaturated substrates, with activities comparable to or exceeding those of related transition-metal-based hydrogenation catalysts. Treatment of the uranium alkyl complex **15** with H₂ results in hydrogenolysis of the uranium–carbon bonds to generate the corresponding uranium dihydride dimer, [(Cp*)₂U(H)]₂(μ -H₂) (**33**), and methane. Hydrogenolysis of this type proceeds through a polar four-centre transition state and features heterolytic H₂ bond scission⁷². No changes in uranium oxidation state are thought to occur despite the fact that higher oxidation states are accessible. The dihydride dimer **33** is unstable in solution and in the solid state, with reductive elimination of H₂ transpiring to form the monohydride dimer, [(Cp*)₂U]₂(μ -H)₂ (**34**) (ref. 4). The dihydride can, however, be stored for extended periods under an atmosphere of H₂, because exposure of **34** to H₂ results in the formation of **33**. Additionally, uranium alkyl complexes may be prepared by 1,2-insertion of alkenes

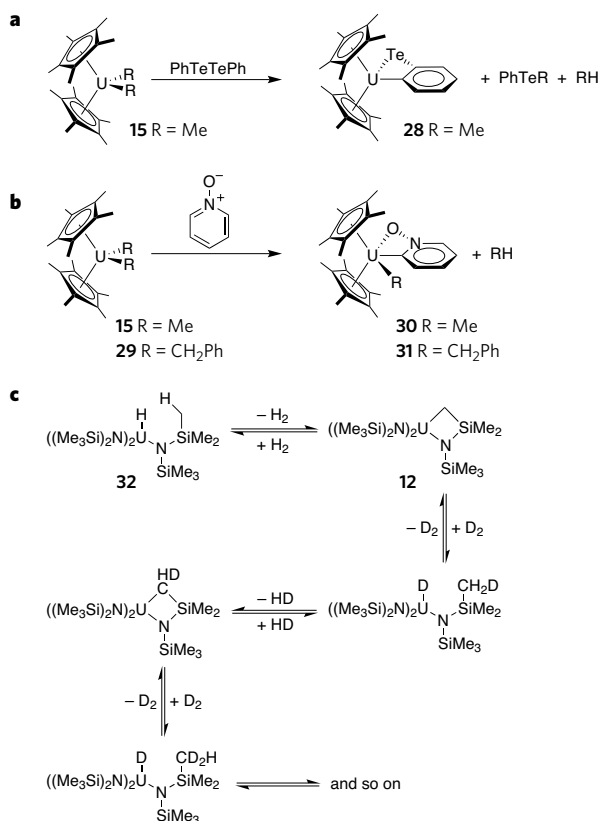


Figure 5 | C–H bond activation with uranium. The direct activation of C–H bonds by uranium offers the opportunity to create value-added organic products. **a**, **b**, Alkane elimination has been coupled with aryl C–H-bond activation by uranium(IV) alkyl complexes. **c**, The reactivity of uranium hydride complex **32** has been shown to include hydrogen–deuterium (H–D) exchange with D₂ by means of a C–H-bond-activating pathway.

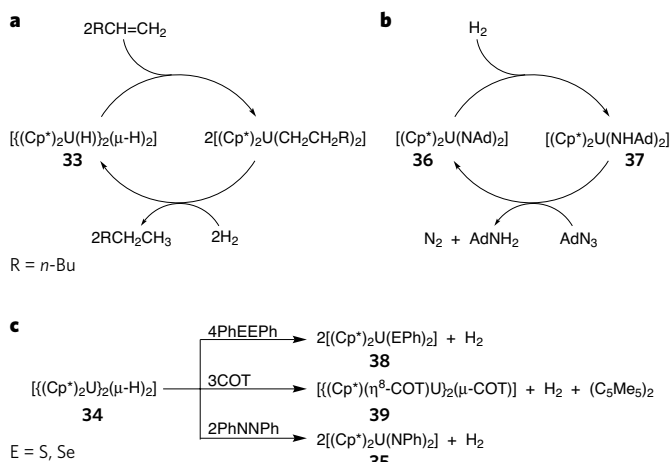


Figure 6 | Hydride and H₂ reactions with uranium. **a**, The uranium dihydride dimer (33) is an active catalyst for the hydrogenation of terminal alkenes. **b**, The uranium bis(imido) complex 36 has been used to effect the catalytic hydrogenation of organic azides by means of a mechanism in which no U–N bond cleavage is thought to occur. **c**, The power and versatility of uranium hydride complexes is no better demonstrated than by the uranium monohydride dimer (34), which performs four-electron, six-electron and even eight-electron reductions of select substrates. COT, 1,3,5,7-cyclo-octatetraene; Cp*, η⁵-C₅Me₅.

into the U–H bond. This reaction has been coupled with hydrogenolysis of the resultant uranium alkyl complex to catalyse the hydrogenation of terminal alkenes and internal alkynes⁴ (Fig. 6a).

In a reaction that is unique to uranium, thermolysis of the uranium(III) hydroxide dimer [$[(\text{Cp}'')_2\text{U}_2(\mu\text{-OH})_2]$ (Cp'' = 1,3-R₂C₅H₃; R = *t*-Bu, SiMe₃) eliminates H₂, generating the uranium(IV) oxo dimer [$[(\text{Cp}'')_2\text{U}_2(\mu\text{-O})_2]$] (ref. 73). Nuclear magnetic resonance and isotopic cross-labelling experiments were consistent with an intramolecular mechanism in which oxidation from U(III/III) to U(IV/IV) occurs along with hydride migration to generate the intermediate, [$[(\text{Cp}'')_2\text{U}(\mu\text{-OH})(\mu\text{-O})\text{U}(\text{H})(\text{Cp}'')_2]$], followed by elimination of H₂ to deliver the observed product. The overall process may be termed 'oxidative elimination' and is a rare counterpoint to the canonical oxidative addition and reductive elimination processes observed across the periodic table.

The catalytic reduction of organic azides and hydrazines to the corresponding amines has been achieved by using [$(\text{Cp}^*)_2\text{U}(\text{NR})_2$] (35, R = Ph; 36, R = Ad) and H₂ (ref. 74) (Fig. 6b). Treatment of 36 with H₂ leads to clean, quantitative reduction to a bis(amido)uranium(IV) complex, [$(\text{Cp}^*)_2\text{U}(\text{NHAd})_2$] (37). In the presence of AdN₃, 37 is oxidized to 36 and both AdNH₂ and N₂ are formed. The catalytic reduction of AdN₃ to AdNH₂ was thus achieved by heating a mixture of 34 and excess AdN₃ under an atmosphere of H₂. When 35 or 36 is treated with *N,N'*-diphenylhydrazine, a 2:1 ratio of aniline and azobenzene are observed as the only organic products, and the starting bis(imido)uranium(vi) complex is completely recovered. Coupled with the absence of any observable intermediates, this suggests a mechanism that lacks U–N bond scission. Rather, the imido ligands mediate H-atom transfer to *N,N'*-diphenylhydrazine while exploiting the U(IV)–U(VI) redox couple.

Liberation of H₂ from uranium hydride complexes has been coupled to the multi-electron reduction of various substrates, in which both metal-based and ligand-based reducing equivalents are used⁷⁵ (Fig. 6c). The nature of the substrate being reduced guides whether oxidation of uranium(III) to uranium(IV) or uranium(VI) occurs. For example, 34 reacts with four equivalents of PhEEP (E = S, Se) to furnish [$(\text{Cp}^*)_2\text{U}(\text{EPH})_2$] (38) and H₂ (one equivalent), an overall four-electron process. The reaction of 34 with three equivalents of COT forms [$(\text{Cp}^*)_2(\eta^8\text{-COT})\text{U}_2(\mu\text{-COT})$] (39), H₂ (one equivalent) and the organic coupling product (C₅Me₅)₂, an overall six-electron process. The reaction of 34 with azobenzene leads to the bis(imido) complex 35 and H₂ (one

equivalent), an overall eight-electron process. In all cases, the same reactivity is observed when 33 is used instead of 34, where initial loss of H₂ from 33 generates 34 before substrate reduction. These results highlight the strongly reducing nature of low-valent uranium complexes, the ability of uranium-bound hydride ligands to be reducing equivalents, and the importance of accessible high-valent states for the resultant uranium complexes.

Reactions of unsaturated groups and polymerizations

Uranium chemistry for both stoichiometric and catalytic reactions of unsaturated functional groups has attained a high level of sophistication in recent years. The kinetic, thermodynamic and mechanistic factors that govern these useful transformations have been extensively reviewed^{10,11}. As researchers continue to refine their understanding of the factors that influence the chemoselectivity and regioselectivity of these reactions, their usefulness for opening routes to valuable organic synthons such as enynes, imines, alkynylimines and vinylsilanes increases.

The dimethyluranium(IV) complex 15 is as a convenient precatalyst for the oligomerization of terminal alkynes. Addition of an alkyne, RC≡CH (two equivalents), leads to rapid loss of methane and formation of the catalytically active bis(acetylide)uranium complex, [$(\text{Cp}^*)_2\text{U}(\text{C}\equiv\text{CR})_2$] (40). In the presence of excess alkyne, 40 catalyses the formation of the geminal 2,4-disubstituted-1-butene-3-yne dimer (41) and the (*E,E*)-1,4,6-trisubstituted-hexa-1,3-diene-5-yne trimer (42) (refs 76, 77). The dimer 41 is generated by 1,2-insertion of alkyne into the U–C bonds in 40, forming the bis(dienyl)uranium intermediate 43. Subsequent σ-bond metathesis with an additional two equivalents of alkyne regenerates 40 and releases 41 (Fig. 7a). The trimer is generated by 2,1-insertion of alkyne into the U–C bonds in 43 to generate intermediate

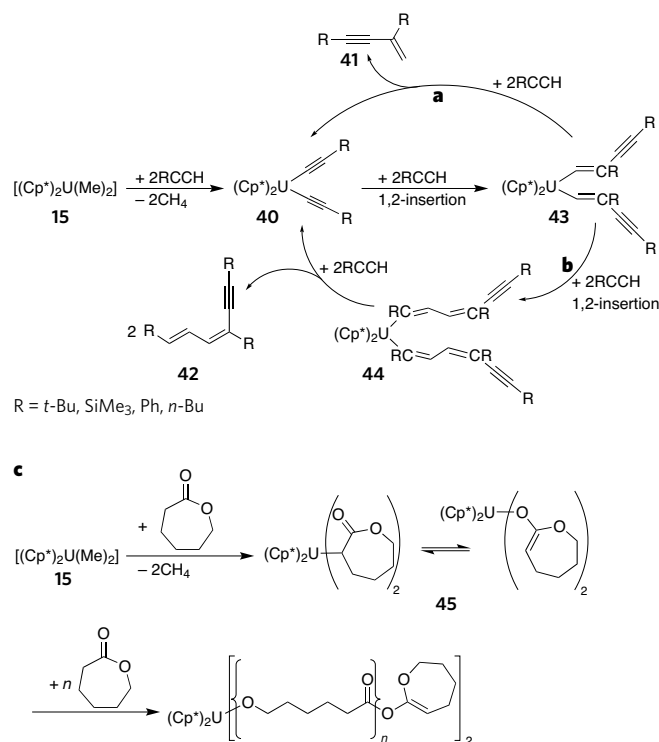


Figure 7 | Oligomerization and polymerization reactions with uranium.

Catalytic oligomerization of alkynes to form enynes 41 and 42 has been achieved with the uranium(IV) alkyl complex (15). **a**, **b**, The reaction proceeds by the insertion of free alkyne into the U–C bond of the catalytically active uranium(IV) bis(acetylide) complex (40). **c**, Complex 15 has also been shown to catalyse the ring-opening polymerization of cyclic esters such as ε-caprolactone. The tautomer-stabilized 45 initiates the polymerization reaction. Notable is the relative facility with which this transformation takes place, despite uranium's oxophilic nature.

44, followed by alkyne σ -bond metathesis to regenerate **40** and release **42** (Fig. 7b). Alkyne cross-oligomerization is achieved when mixtures of alkynes are used⁷⁸. In either case, the regioselectivity of the reaction is governed by steric interactions between the substituent on the incoming alkyne and the ancillary ligand framework or the dienyl fragment. The triamidouranium(IV) complex $[(Et_2N)_3U][BPh_4]$ also catalyses the oligomerization of terminal alkenes to the geminal enyne dimer and higher oligomers⁷⁹. The regioselectivity profile of the reaction is reversed when performed in the presence of primary amines, $R'NH_2$. The product oligomer distribution for the reaction varies depending on the nature of R and R' , and in certain cases is highly selective for *cis*-enyne dimer formation⁸⁰.

The coupling of terminal alkynes and isonitriles ($RN\equiv C$) is catalysed by **15**, generating an (*E*)-alkynylimine by 1,1-insertion of isonitrile into the uranium–acetylide bond of catalytically active **40** (ref. 81). Higher oligomers are also obtained, and the product distributions have been shown to be highly dependent on alkyne and isonitrile concentrations. The ion pair $[(Et_2N)_3U][BPh_4]$ also catalyses the coupling of alkynes and isonitriles, providing exclusively the (*E*)-alkynylimine product⁸¹.

A diuranium *ansa*-metallocene complex featuring a bridging oxide ligand has been shown to catalyse the metathesis of silylalkynes and the cross-metathesis of silylalkynes with terminal alkynes⁸². In the proposed mechanism, a uranium silylacetylide complex, $[(Me_2Si(\eta^5-C_5Me_5)_2)U(C\equiv CSiMe_3)_2(\mu-O)]$, reacts with $Me_3SiC\equiv CH$ to form $Me_3SiC\equiv CSiMe_3$ and the parent uranium acetylide derivative, $[(Me_2Si(\eta^5-C_5Me_5)_2)U(C\equiv CH)_2(\mu-O)]$. Protonolysis with an additional equivalent of $Me_3SiC\equiv CH$ eliminates ethyne and $[(Me_2Si(\eta^5-C_5Me_5)_2)U(C\equiv CSiMe_3)_2(\mu-O)]$ is reformed. The critical transformation involves a four-centre transition state and exchanges Si–C and U–C σ -bonds. This mechanism differs significantly from that of alkyne metathesis with certain transition-metal alkylidyne complexes, in which the alkyne C–C triple bond participates directly in a four-centre metalocyclic intermediate⁸³.

Both **15** and $[(Et_2N)_3U][BPh_4]$ have been shown to be active for the ring-opening polymerization of ϵ -caprolactone and L-lactide⁸⁴. For **15**, σ -bond activation leads to the tautomer-stabilized uranium enolate complex, **45**, which reacts with excess monomer to provide polymers with high molecular masses and low polydispersity (Fig. 7c). In addition to demonstrating that industrially relevant polyester syntheses can be catalysed by uranium complexes, these results most importantly show that the oxophilic nature of uranium is not necessarily an insurmountable thermodynamic barrier to achieving turnover in catalytic reactions involving intermediate U–O bonds.

Hydroamination and hydrosilation

The activation of amine N–H bonds with uranium complexes leads to a range of uranium imido and uranium amido complexes relevant to hydroamination catalysis and beyond. The reaction of $[(Cp'')_2U(Me)_2]$ (**46**) with primary amines results in the elimination of methane and formation of the bis(amido) complexes $[(Cp'')_2U(NHR)_2]$ (**47**, $R = Me$; **48**, $R = p$ -tolyl; **49**, $R = CH_2Ph$)⁸⁵. At elevated temperatures, **47** and **48** enter an equilibrium with the monomeric uranium(IV) imido complex, $[(Cp'')_2U(NR)]$ (**50**, $R = Me$; **51**, $R = p$ -tolyl), by eliminating an equivalent of the corresponding primary amine. The less sterically crowded **15** has similar reactivity; however, only with larger substituents at the amine does the equilibrium between bis(amido) and imido complexes become apparent⁸⁶. Activation of ammonia has been observed with $[(\eta^5-1,3-(Me_3E)_2C_5H_3)_2U(Me)_2]$ ($E = C, Si$), a reaction resulting in the bridging uranium(IV) bis(imido) species $[(\eta^5-1,3-(Me_3E)_2C_5H_3)_2U]_2(\mu-NH_2)_2]$ (ref. 85).

Hydroamination of alkynes has been well demonstrated with a variety of uranium complexes¹¹. For example, $[(Cp^*)_2U(NHR)_2]$ (**52**) catalyses the hydroamination of terminal alkynes with primary amines, chemoselectively and regioselectively forming the (*E*)-imine product, **53** (ref. 87) (Fig. 8b). The mechanism is thought to involve an intermediate imidouranium(IV) complex, $[(Cp^*)_2U(NR)]$ (**54**), in which the imido ligand is derived from the primary amine as described

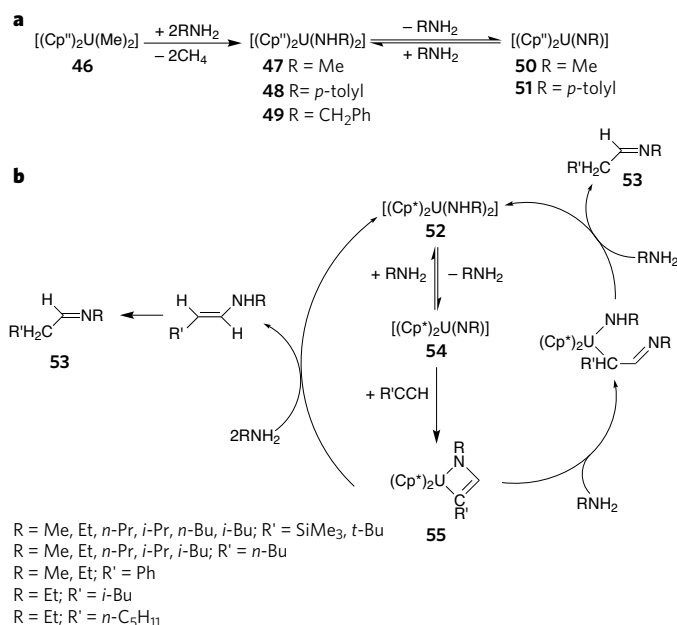


Figure 8 | Hydroamination of alkynes with uranium. **a**, Bis(amido) complexes of uranium(IV) can function as precursors to the corresponding uranium(IV) imido complexes by elimination of the corresponding primary amine. **b**, This feature has been exploited to catalyse the hydroamination of alkynes, in which the elimination of amine from the bis(amido) complex **52** leads to the imido product **54**. Alkynes engage **54** by cycloaddition to form the metalocyclobutene intermediate **55**, which further reacts with free amine to exclusively form an *E*-imine product (**53**) while regenerating **52**. Cp'', 1,3-R₂C₅H₃; Et, ethyl; *i*-Bu, isobutyl.

above. Subsequent 2,1-insertion of alkyne into the U–N bond leads to an azametallocyclobutene intermediate (**55**), which reacts further with free amine to release the organic product and regenerate **52**. This mechanism is analogous to that described for the hydroamination of alkynes with $[(Cp)_2Zr(NHR)_2]$ (ref. 88). Hydroamination of terminal alkynes is also accomplished by using $[(Et_2N)_3U][BPh_4]$ as a precatalyst⁸⁰. Varying the amines in this case affects the stereochemistry of the product, and more bulky amines provide a mixture. The constrained-geometry catalyst, $(CGC)U(NRR')_2$ ($CGC = Me_2Si(\eta^5-Me_4C_5)(t-Bu)N$; $R, R' = Me, Et$), is an efficient precatalyst for the catalytic intramolecular hydroamination and cyclization of aminoalkenes, aminoalkynes, aminoallenes and aminodienes^{89,90}. The catalyst's efficiency is attributed to the steric accessibility of the uranium centre as enforced by the constrained geometry; the catalyst is stabilized throughout the reaction by the coordination of free amine.

Hydrosilation of terminal alkynes with phenylsilane has been achieved with **15** as a precatalyst⁹¹. The product distribution contains the *trans*-vinylsilane $RCH=CHSiH_2Ph$ ($R = i$ -Pr, n -Bu, t -Bu, $SiMe_3$) as the major component, with the alkene $RCH=CH_2$ and dehydrocoupled silylalkyne $RC\equiv CSiH_2Ph$ as minor components. The ion pair $[(Et_2N)_3U][BPh_4]$ catalyses the hydrosilation of terminal alkynes as well as the dehydrocoupling of amines and silanes⁹².

Conclusions

The stoichiometric transformations presented here satisfy the first two criteria for catalysis: the binding and subsequent reaction of substrate. The third criterion, release of the newly functionalized product, presents the greatest challenge to the expanding field of catalysis with uranium. Successfully turning over a catalytic cycle requires a careful balance of bond dissociation enthalpies among participants along the reaction coordinate. The tendency of uranium ions to form very strong bonds to hard, anionic ligands, just the sort of ligands that are formed in the reductive activation of small molecules, often results in a thermodynamic well from which extricating the functionalized product and regenerating the uranium catalyst are difficult. This problem may

be circumvented in several ways. Ancillary ligands designed to inhibit optimal coordination of substrate to uranium can help to destabilize intermediate complexes, thus allowing the product to disengage from the uranium centre more readily. Another approach casts uranium as a stoichiometric participant in a domino catalysis reaction, by which a non-uranium-based catalyst effects the release of the functionalized substrate and returns the uranium complex to its initial reactive form⁹³. Although not catalysis with uranium in the strictest sense, this approach would allow the unique stoichiometric reactivity of uranium to be incorporated into overall catalytic processes. Indeed, the possibilities offered by using uranium complexes in concert with other metals to activate otherwise unreactive uranium–ligand bonds were recently demonstrated by Polly Arnold and co-workers⁹⁴: a uranyl complex was coaxed into cleaving C–Si and N–Si bonds by polarizing the [UO₂]²⁺ core with coordinated potassium cations.

The study of catalysis with transition-metal-based systems is by comparison a mature field, one from which many benefits have been derived. Research on uranium catalysis, although much less developed, has already led to application in industrial settings⁹⁵. In part, a reluctance to embrace uranium for scientific and technological applications can be traced to anti-nuclear public sentiment; to this objection we respond that depleted uranium is only weakly radioactive and can be manipulated safely in macroscopic quantities by trained individuals. Ultimately, establishing the utility of uranium complexes for common synthetic applications depends on demonstrating that uranium catalysts offer practical advantages not found among the transition metals. Compelling arguments for the application of uranium-based catalysts may lie in the functionalization of inert small molecules such as N₂, CO and CO₂ under mild conditions. This chemistry lies at the very forefront of the field, and many questions about the reactivity of these and other uranium complexes remain to be answered. Only by charting the reactivity of known molecular uranium systems, and by creating new ones, can chemists identify ideally suited applications in catalysis for this unique and abundant element. In the course of this endeavour we may contemplate a satisfying transformation: that of waste product to valuable catalyst. ■

Note added in proof: Multiple bond metathesis chemistry with CO₂ has also been observed for uranium(v) imido complexes with the [(^kArO)₃tacn)] ligand set, forming the uranium(v) terminal oxo species via extrusion of free isocyanate⁹⁶. Using such high-valent uranium precursors for the synthesis of terminal oxo functionalities eliminates the decomposition and dimerization reactions observed when the previously reported low-valent uranium(III) analogues were treated with CO₂ or traditional oxygen-atom transfer reagents.

- Bart, S. C. & Meyer, K. in *Structure and Bonding* Vol. 127 (ed. Albrecht-Schmidt, T. E.) 119–176 (Springer, 2008).
This book chapter provides a detailed discussion of recent developments in the area of coordination chemistry and small-molecule activation with uranium.
- Ephritikhine, M. The vitality of uranium molecular chemistry at the dawn of the XXIst century. *Dalton Trans.* 2501–2516 (2006).
This paper provides an enlightening perspective on the state of synthetic uranium chemistry at the beginning of the twenty-first century.
- Burns, C. J. & Eisen, M. S. in *The Chemistry of the Actinide and Transactinide Elements* Vol. 5 (eds Morss, L. R., Edelstein, N. M. & Fuge, J.) 2911–3012 (Springer, 2006).
This book chapter is a detailed review of catalysis with uranium.
- Fagan, P. J., Manriquez, J. M., Maatta, E. A., Seyam, A. M. & Marks, T. J. Synthesis and properties of bis(pentamethylcyclopentadienyl) actinide hydrocarbyls and hydrides. A new class of highly reactive f-element organometallic compounds. *J. Am. Chem. Soc.* **103**, 6650–6667 (1981).
- Bruno, J. W., Duttera, M. R., Fendrick, C. M., Smith, G. M. & Marks, T. J. New stoichiometric and catalytic organometallic chemistry with actinides. C–H activation and phosphine/phosphite coordination chemistry. *Inorg. Chim. Acta* **94**, 271–277 (1984).
- He, M. Y., Xiong, G., Toscano, P. J., Burwell, R. L. & Marks, T. J. Supported organoactinides. Surface chemistry and catalytic properties of alumina-bound cyclopentadienyl and penta methylcyclopentadienyl thorium and uranium hydrocarbyls and hydrides. *J. Am. Chem. Soc.* **107**, 641–652 (1985).
- Nolan, S. P., Porchia, M. & Marks, T. J. Organo-f-element thermochemistry. Actinide-group 14 element and actinide-transition-element bond disruption enthalpies and stoichiometric/catalytic chemical implications thereof in heterobimetallic tris(cyclopentadienyl)uranium(IV) compounds. *Organometallics* **10**, 1450–1457 (1991).
- Lin, Z. & Marks, T. J. A kinetic, mechanistic, and molecular mechanics investigation of olefin insertion into organoactinide-hydride bonds. Metal, olefin, ancillary ligand, and

- diastereoselection effects. *J. Am. Chem. Soc.* **112**, 5515–5525 (1990).
- Ephritikhine, M. Synthesis, structure, and reactions of hydride, borohydride, and aluminohydride compounds of the f-elements. *Chem. Rev.* **97**, 2193–2242 (1997).
This paper contains useful discussions of uranium hydride complexes and the chemistry that they engage in.
- Barnea, E. & Eisen, M. S. Organoactinides in catalysis. *Coord. Chem. Rev.* **250**, 855–899 (2006).
This review provides a detailed discussion of catalysis with uranium, emphasizing the important transformations of alkyne substrates, such as oligomerization, hydroamination and hydrosilation.
- Andrea, T. & Eisen, M. S. Recent advances in organothorium and organouranium catalysis. *Chem. Soc. Rev.* **37**, 550–567 (2008).
- Haber, F. Ammonia. German patent DE 229126 (1909).
- Roussel, P. & Scott, P. Complex of dinitrogen with trivalent uranium. *J. Am. Chem. Soc.* **120**, 1070–1071 (1998).
- Cloke, F. G. N. & Hitchcock, P. B. Reversible binding and reduction of dinitrogen by a uranium(III) pentalene complex. *J. Am. Chem. Soc.* **124**, 9352–9353 (2002).
- Odom, A. L., Arnold, P. L. & Cummins, C. C. Heterodinuclear uranium/molybdenum dinitrogen complexes. *J. Am. Chem. Soc.* **120**, 5836–5837 (1998).
- Eikey, R. A. & Abu-Omar, M. M. Nitrido and imido transition metal complexes of groups 6–8. *Coord. Chem. Rev.* **243**, 83–124 (2003).
- Korobkov, I., Gambarotta, S. & Yap, G. P. A. A highly reactive uranium complex supported by the calix[4]tetrapyrrole tetraanion affording dinitrogen cleavage, solvent deoxygenation, and polysilanol depolymerization. *Angew. Chem. Int. Edn Engl.* **41**, 3433–3436 (2002).
- MacKay, B. A. & Fryzuk, M. D. Dinitrogen coordination chemistry: on the biomimetic borderlands. *Chem. Rev.* **104**, 385–401 (2004).
- Pool, J. A., Bernskoetter, W. H. & Chirik, P. J. On the origin of dinitrogen hydrogenation promoted by [(η⁵-C₅Me₅H)₂Zr]₂(μ²-η², η²-N₂). *J. Am. Chem. Soc.* **126**, 14326–14327 (2004).
- Chirik, P. J. Dinitrogen functionalization with bis(cyclopentadienyl) complexes of zirconium and hafnium. *Dalton Trans.* 16–25 (2007).
- Schrock, R. R. Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. *Acc. Chem. Res.* **38**, 955–962 (2005).
- Brennan, J. G., Andersen, R. A. & Robbins, J. L. Preparation of the first molecular carbon monoxide complex of uranium, (Me₃SiC₅H₄)₃UO. *J. Am. Chem. Soc.* **108**, 335–336 (1986).
- Parry, J., Carmona, E., Coles, S. & Hursthouse, M. Synthesis and single crystal X-ray diffraction study on the first isolable carbonyl complex of an actinide, (C₅Me₅H)₃U(CO). *J. Am. Chem. Soc.* **117**, 2649–2650 (1995).
- Sikora, D. J., Rausch, M. D., Rogers, R. D. & Atwood, J. L. New syntheses and molecular structures of the decamethylmetallocene dicarbonyls (η⁵-C₅Me₅)₂M(CO)₂ (M = titanium, zirconium, hafnium). *J. Am. Chem. Soc.* **103**, 1265–1267 (1981).
- Evans, W. J., Kozimor, S. A., Nyce, G. W. & Ziller, J. W. Comparative reactivity of sterically crowded η² (C₅Me₅)₃Nd and (C₅Me₅)₃U complexes with CO: formation of a nonclassical carbonium ion versus an element metal carbonyl complex. *J. Am. Chem. Soc.* **125**, 13831–13835 (2003).
- Cotton, F. A., Wilkinson, G., Murillo, C. A. & Bochmann, M. *Advanced Inorganic Chemistry*. 635–639 (Wiley, 1999).
- Castro-Rodriguez, I. & Meyer, K. Carbon dioxide reduction and carbon monoxide activation employing a reactive uranium(III) complex. *J. Am. Chem. Soc.* **127**, 11242–11243 (2005).
- Summerscales, O. T., Cloke, F. G. N., Hitchcock, P. B., Green, J. C. & Hazari, N. Reductive cyclotrimerization of carbon monoxide to the deltatate dianion by an organometallic uranium complex. *Science* **311**, 829–831 (2006).
- Summerscales, O. T., Cloke, F. G. N., Hitchcock, P. B., Green, J. C. & Hazari, N. Reductive cyclotetramerization of CO to squarate by a U(III) complex: the X-ray crystal structure of [U(η⁵-C₅H₅SiPr₂-1,4)₂](η⁵-C₅Me₅H)₂(μ²-η²-C₄O₄). *J. Am. Chem. Soc.* **128**, 9602–9603 (2006).
This paper describes an excellent example of CO activation with uranium and provides useful leading references.
- Weydert, M., Brennan, J. G., Andersen, R. A. & Bergman, R. G. Reactions of a uranium(IV) tertiary alkyl bond: facile ligand-assisted reduction and insertion of ethylene and carbon monoxide. *Organometallics* **14**, 3942–3951 (1995).
- Paolucci, G., Rossetto, G., Zanella, P., Yunlu, K. & Fischer, R. D. Facile insertion of carbon monoxide into the uranium–carbon and uranium–nitrogen bonds of some tris(cyclopentadienyl)uranium(IV)-alkyl and -dialkylamide derivatives. *J. Organometall. Chem.* **272**, 363–383 (1984).
- Dormond, A., Aaliti, A., Elbouadili, A. & Moise, C. Reactivity of uranium–carbon σ bonds: reactions of methyltris(hexamethyldisilylamido)uranium. *J. Organometall. Chem.* **329**, 187–199 (1987).
- Simpson, S. J. & Andersen, R. A. Actinide–carbon bonds: insertion reactions of carbon monoxide, *tert*-butyl isocyanide, and *tert*-butyl cyanide into [(Me₃Si)₂N]₂MCH₂Si(Me)₂NSiMe₃. *J. Am. Chem. Soc.* **103**, 4063–4066 (1981).
- Cramer, R. E., Maynard, R. B., Paw, J. C. & Gilje, J. W. Carbon monoxide insertion into a uranium carbon double bond. The structure of (η⁵-C₅H₅)₂U(η²-OCCH)(P)(CH₃)(C₆H₅)₂. *Organometallics* **1**, 869–871 (1982).
- Louie, J., Gibby, J. E., Farnworth, M. V. & Tekavec, T. N. Efficient nickel-catalyzed [2 + 2 + 2] cycloaddition of CO₂ and diynes. *J. Am. Chem. Soc.* **124**, 15188–15189 (2002).
- Takimoto, M. & Mori, M. Novel catalytic CO₂ incorporation reaction: nickel-catalyzed regio- and stereoselective ring-closing carboxylation of bis-1,3-dienes. *J. Am. Chem. Soc.* **124**, 10008–10009 (2002).
- Calderazzo, F., Dell'Amico, G., Netti, R. & Pasquali, M. Dialkylcarbamato complexes of transition elements. I. A new method for the synthesis of *N,N*-dialkylcarbamato and *N,N*-dialkylthiocarbamato complexes of uranium(IV). *Inorg. Chim. Acta* **17**, 471–473 (1978).
- Lescop, C., Arliguie, T., Lance, M., Nierlich, M. & Ephritikhine, M. Bis(pentamethylcyclopentadienyl)uranium(IV) thiolate compounds. Synthesis and reactions with CO₂ and CS₂. *J. Organometall. Chem.* **580**, 137–144 (1999).
- Moloy, K. G. & Marks, T. J. The insertion of carbon-dioxide into actinide alkyl and hydride bonds. *Inorg. Chim. Acta* **110**, 127–131 (1985).
- Castro-Rodriguez, I., Nakai, H., Zakharov, L. N., Rheingold, A. L. & Meyer, K. A linear,

- O-coordinated η^1 -CO₂ bound to uranium. *Science* **305**, 1757–1760 (2004).
41. Harris, D. C. & Bertolucci, M. D. *Symmetry and Spectroscopy: An Introduction to Vibrational and Electronic Spectroscopy* (Dover, 1978).
 42. Mindiola, D. J. Early transition-metal hydrazido complexes: masked metallanitrenes from N–N bond scission. *Angew. Chem. Int. Edn Engl.* **47**, 1557–1559 (2008).
 43. Thetford, R. & Mignanelli, M. The chemistry and physics of modelling nitride fuels for transmutation. *J. Nucl. Mater.* **320**, 44–53 (2003).
 44. Evans, W. J., Kozimor, S. A. & Ziller, J. W. Molecular octa-uranium rings with alternating nitride and azide bridges. *Science* **309**, 1835–1838 (2005).
 45. Nocton, G., Pécaut, J. & Mazzanti, M. A nitrido-centered uranium azido cluster obtained from a uranium azide. *Angew. Chem. Int. Edn Engl.* **47**, 3040–3042 (2008).
 46. Crawford, M.-J., Ellern, A. & Mayer, P. UN₃³⁻: a structurally characterized binary actinide heptaazide anion. *Angew. Chem. Int. Edn Engl.* **44**, 7874–7878 (2005).
 47. Zi, G. *et al.* Preparation and reactions of base-free bis(1,2,4-tri-*tert*-butylcyclopentadienyl) uranium oxide, Cp⁺UO. *Organometallics* **24**, 4251–4264 (2005).
 48. Brennan, J. G. & Andersen, R. A. Electron-transfer reactions of trivalent uranium. Preparation and structure of the uranium metallocene compounds (MeC₅H₄)₂U–Nph and [(MeC₅H₄)₂U]₂[μ-η¹,η²-PhNCO]. *J. Am. Chem. Soc.* **107**, 514–516 (1985).
 49. Arney, D. S. J., Burns, C. J. & Smith, D. C. Synthesis and structure of the first uranium(vi) organometallic complex. *J. Am. Chem. Soc.* **114**, 10068–10069 (1992).
 50. Arney, D. S. J. & Burns, C. J. Synthesis and properties of high-valent organouranium complexes containing terminal organoimido and oxo functional groups. A new class of organo-f-element complexes. *J. Am. Chem. Soc.* **117**, 9448–9460 (1995).
 51. Warner, B. P., Scott, B. L. & Burns, C. J. A simple preparative route to bis(imido)uranium(vi) complexes by the direct reductions of diazenes and azides. *Angew. Chem. Int. Edn Engl.* **37**, 959–960 (1998).
 52. Roussel, P., Boaretto, R., Kingsley, A. J., Alcock, N. W. & Scott, P. Reactivity of a triamidoamine complex of trivalent uranium. *J. Chem. Soc. Dalton Trans.* 1423–1428 (2002).
 53. Zalkin, A., Brennan, J. G. & Andersen, R. A. Tris[bis(trimethylsilyl)amido](trimethylsilylimido)uranium(v). *Acta Crystallogr. C* **44**, 1553–1554 (1988).
 54. Castro-Rodríguez, I., Olsen, K., Gantzel, P. & Meyer, K. Uranium tris-aryloxide derivatives supported by triazacyclononane: engendering a reactive uranium(III) center with a single pocket for reactivity. *J. Am. Chem. Soc.* **125**, 4565–4571 (2003).
 55. Castro-Rodríguez, I., Nakai, H. & Meyer, K. Multiple-bond metathesis mediated by sterically pressured uranium complexes. *Angew. Chem. Int. Edn Engl.* **45**, 2389–2392 (2006).
- This paper describes an interesting synthetic cycle with uranium and paints a picture of the possibilities for catalysis contained in such an approach.**
56. Hayton, T. W. *et al.* Synthesis of imido analogs of the uranyl ion. *Science* **310**, 1941–1943 (2005).
 57. Hayton, T. W., Boncella, J. M., Scott, B. L., Batista, E. R. & Hay, P. J. Synthesis and reactivity of the imido analogues of the uranyl ion. *J. Am. Chem. Soc.* **128**, 10549–10559 (2006).
 58. Hayton, T. W., Boncella, J. M., Scott, B. L. & Batista, E. R. Exchange of an imido ligand in bis(imido) complexes of uranium. *J. Am. Chem. Soc.* **128**, 12622–12623 (2006).
 59. Spencer, L. P., Yang, P., Scott, B. L., Batista, E. R. & Boncella, J. M. Imido exchange in bis(imido) uranium(vi) complexes with aryl isocyanates. *J. Am. Chem. Soc.* **130**, 2930–2931 (2008).
- This paper highlights new developments in the chemistry of uranyl analogues and contains useful leading references.**
60. Smith, G. M., Carpenter, J. D. & Marks, T. J. Intramolecular vs. intermolecular alkyl carbon–hydrogen bond activation. Complete thermodynamic and kinetic parameters for a reversible cyclometalation. *J. Am. Chem. Soc.* **108**, 6805–6807 (1986).
 61. Ciliberto, E. *et al.* Combined DV-Xα and gas-phase UV photoelectron spectroscopic investigation of the electronic structures of tetravalent titanium, zirconium, molybdenum, and thorium 1-sila-3-metallacyclobutane metallocene complexes. *Organometallics* **11**, 1727–1737 (1992).
 62. Evans, W. J. *et al.* Formation of (C₅Me₃)₂U(EPh)Me, (C₅Me₃)₂U(EPh)₂, and (C₅Me₃)₂U(η²-TeC₆H₄) from (C₅Me₃)₂UMe₂ and PhEPh (E = S, Se, Te). *Organometallics* **26**, 4287–4293 (2007).
 63. Kiplinger, J. L., Morris, D. E., Scott, B. L. & Burns, C. J. Convenient synthesis, structure, and reactivity of (C₅Me₃)₂U(CH₂C₆H₅)₃: a simple strategy for the preparation of monopentamethylcyclopentadienyl uranium(IV) complexes. *Organometallics* **21**, 5978–5982 (2002).
 64. Pool, J. A., Scott, B. L. & Kiplinger, J. L. A new mode of reactivity for pyridine N-oxide: C–H activation with uranium(IV) and thorium(IV) bis(alkyl) complexes. *J. Am. Chem. Soc.* **127**, 1338–1339 (2005).
 65. Pool, J. A., Scott, B. L. & Kiplinger, J. L. Synthesis of actinide η²-pyridyl and η²-α-picoyl complexes by carbon–hydrogen bond activation. *J. Alloys Compd.* **418**, 178–183 (2006).
 66. Kiplinger, J. L., Scott, B. L., Schelter, E. J. & Pool Davis Tourneay, J. A. sp² versus sp² C–H bond activation chemistry of 2-picoline by Th(IV) and U(IV) metallocene complexes. *J. Alloys Compd.* **444–445**, 477–482 (2007).
 67. Lam, O. P., Feng, P. L., Heinemann, F. W., O'Connor, J. M. & Meyer, K. Charge-separation in uranium diazomethane complexes leading to C–H activation and chemical transformation. *J. Am. Chem. Soc.* **130**, 2806–2816 (2008).
 68. Simpson, S. J., Turner, H. W. & Andersen, R. A. Hydrogen–deuterium exchange: perdeuteriohydrottris(hexamethyldisilylamido)thorium(IV) and -uranium(IV). *J. Am. Chem. Soc.* **101**, 7728–7729 (1979).
 69. Simpson, S. J., Turner, H. W. & Andersen, R. A. Preparation and hydrogen–deuterium exchange of alkyl and hydride bis(trimethylsilyl)amido derivatives of the actinide elements. *Inorg. Chem.* **20**, 2991–2995 (1981).
 70. Castro-Rodríguez, I. *et al.* Evidence for alkane coordination to an electron-rich uranium center. *J. Am. Chem. Soc.* **125**, 15734–15735 (2003).
 71. Deacon, G. B., Tuong, T. D. & Franklin, G. R. Reactions of uranium hydride with some organic and organometallic compounds. *Inorg. Chim. Acta* **111**, 193–196 (1986).
 72. Lin, Z. & Marks, T. J. Metal, bond energy, and ancillary ligand effects on actinide–carbon σ-bond hydrogenolysis. A kinetic and mechanistic study. *J. Am. Chem. Soc.* **109**, 7979–7985 (1987).
 73. Lukens, W. W., Beshouri, S. M., Blosch, L. L. & Andersen, R. A. Oxidative elimination of H₂ from [Cp⁺₂U(μ-OH)]₂ to form [Cp⁺₂U(μ-O)]₂, where Cp⁺ is 1,3-(Me₃C)₂C₅H₃ or 1,3-(Me₃Si)₂C₅H₃. *J. Am. Chem. Soc.* **118**, 901–902 (1996).
 74. Peters, R. G., Warner, B. P. & Burns, C. J. The catalytic reduction of azides and hydrazines using high-valent organouranium complexes. *J. Am. Chem. Soc.* **121**, 5585–5586 (1999).
 75. Evans, W. J. *et al.* Actinide hydride complexes as multielectron reductants: analogous reduction chemistry from [(C₅Me₃)₂UH]₂, [(C₅Me₃)₂UH]₂, and [(C₅Me₃)₂ThH]₂. *Organometallics* **26**, 3568–3576 (2007).
- This paper highlights the versatility of uranium hydrides for effecting multi-electron reductions.**
76. Straub, T., Haskel, A. & Eisen, M. S. Organoactinide-catalyzed oligomerization of terminal acetylenes. *J. Am. Chem. Soc.* **117**, 6364–6365 (1995).
 77. Haskel, A., Wang, J. Q., Straub, T., Neyroud, T. G. & Eisen, M. S. Controlling the catalytic oligomerization of terminal alkynes promoted by organoactinides: a strategy to short oligomers. *J. Am. Chem. Soc.* **121**, 3025–3034 (1999).
 78. Haskel, A., Straub, T., Dash, A. K. & Eisen, M. S. Oligomerization and cross-oligomerization of terminal alkynes catalyzed by organoactinide complexes. *J. Am. Chem. Soc.* **121**, 3014–3024 (1999).
 79. Wang, J. Q., Dash, A. K., Berthet, J. C., Ephritikhine, M. & Eisen, M. S. Selective dimerization of terminal alkynes promoted by the cationic actinide compound [(Et₂N)₃U][BPh₄]. Formation of the alkyne π-complex [(Et₂N)₂U(C≡C'Bu)(η¹-HC≡C'Bu)][BPh₄]. *Organometallics* **18**, 2407–2409 (1999).
 80. Wang, J. *et al.* Oligomerization and hydroamination of terminal alkynes promoted by the cationic organoactinide compound [(Et₂N)₃U][BPh₄]. *Chem. Eur. J.* **8**, 5384–5396 (2002).
 81. Barnea, E. *et al.* Catalytic coupling of terminal alkynes with isonitriles promoted by organoactinide complexes. *J. Am. Chem. Soc.* **126**, 10860–10861 (2004).
 82. Wang, J., Gurevich, Y., Botoshansky, M. & Eisen, M. S. Unique σ-bond metathesis of silylalkynes promoted by an *ansa*-dimethylsilyl and oxo-bridged uranium metallocene. *J. Am. Chem. Soc.* **128**, 9350–9351 (2006).
 83. Schrock, R. R. High-oxidation-state molybdenum and tungsten alkylidyne complexes. *Acc. Chem. Res.* **19**, 342–348 (1986).
 84. Barnea, E., Moradove, D., Berthet, J. C., Ephritikhine, M. & Eisen, M. S. Surprising activity of organoactinide complexes in the polymerization of cyclic mono- and diesters. *Organometallics* **25**, 320–322 (2006).
 85. Zi, G., Blosch, L. L., Jia, L. & Andersen, R. A. Preparation and reactions of base-free bis(1,2,4-tri-*tert*-butylcyclopentadienyl)uranium methylimide, Cp⁺U=NMe, and related compounds. *Organometallics* **24**, 4602–4612 (2005).
 86. Straub, T., Frank, W., Reiss, G. J. & Eisen, M. S. Uranium(IV) bis(amido), imido and bis(acetylide) complexes: synthesis, molecular structure, solution dynamics and interconversion reactions. *J. Chem. Soc. Dalton Trans.* 2541–2546 (1996).
 87. Haskel, A., Straub, T. & Eisen, M. S. Organoactinide-catalyzed intermolecular hydroamination of terminal alkynes. *Organometallics* **15**, 3773–3775 (1996).
 88. Pohki, F. & Doye, S. The catalytic hydroamination of alkynes. *Chem. Soc. Rev.* **32**, 104–114 (2003).
 89. Stubbert, B. D., Stern, C. L. & Marks, T. J. Synthesis and catalytic characteristics of novel constrained-geometry organoactinide catalysts. The first example of actinide-mediated intramolecular hydroamination. *Organometallics* **22**, 4836–4838 (2003).
 90. Stubbert, B. D. & Marks, T. J. Constrained geometry organoactinides as versatile catalysts for the intramolecular hydroamination/cyclization of primary and secondary amines having diverse tethered C–C unsaturation. *J. Am. Chem. Soc.* **129**, 4253–4271 (2007).
 91. Dash, A. K., Wang, J. Q. & Eisen, M. S. Catalytic hydrosilylation of terminal alkynes promoted by organoactinides. *Organometallics* **18**, 4724–4741 (1999).
 92. Dash, A. K., Wang, J. X., Berthet, J. C., Ephritikhine, M. & Eisen, M. S. Diverse catalytic activity of the cationic actinide complex [(Et₂N)₃U][BPh₄] in the dimerization and hydrosilylation of terminal alkynes. Characterization of the first f-element alkyne π-complex [(Et₂N)₂U(C≡C'Bu)(η¹-HC≡C'Bu)][BPh₄]. *J. Organometall. Chem.* **604**, 83–98 (2000).
 93. Fogg, D. E. & dos Santos, E. N. Tandem catalysis: a taxonomy and illustrative review. *Coord. Chem. Rev.* **248**, 2365–2379 (2004).
 94. Arnold, P. L., Patel, D., Wilson, C. & Love, J. B. Reduction and selective oxo group silylation of the uranyl dication. *Nature* **451**, 315–317 (2008).
- This paper describes a recent and important development in the chemistry of uranyl compounds.**
95. Cottingham, P. L. & Barker, L. K. Depleted uranium as catalyst for hydrocracking shale oil. *Ind. Eng. Chem. Prod. Res. Dev.* **12**, 41–47 (1973).
 96. Bart, S. C. *et al.* Carbon dioxide activation with sterically pressured mid- and high-valent uranium complexes. *J. Am. Chem. Soc.* doi:10.1021/ja804263w (in the press).

Acknowledgements We thank the National Science Foundation (grant number CHE-0724158), the US Department of Energy (grant number DE-FG02-O4ER15537), the German Research Foundation (DFG; grant number ME-1754/2-1) and the Sonderforschungsbereich (grant number 583) for financial support. A.R.F. acknowledges Massachusetts Institute of Technology for the Alan Davison Graduate Fellowship. S.C.B. acknowledges the Alexander von Humboldt Foundation for a postdoctoral fellowship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence should be addressed to C.C.C. (ccummins@mit.edu).

An intrinsic mechanism of corticogenesis from embryonic stem cells

Nicolas Gaspard¹, Tristan Bouchet¹, Raphael Hourez², Jordane Dimidschstein¹, Gilles Naeije¹, Jelle van den Aamee¹, Ira Espuny-Camacho¹, Adèle Herpoel¹, Lara Passante¹, Serge N. Schiffmann², Afsaneh Gaillard³ & Pierre Vanderhaeghen¹

The cerebral cortex develops through the coordinated generation of dozens of neuronal subtypes, but the mechanisms involved remain unclear. Here we show that mouse embryonic stem cells, cultured without any morphogen but in the presence of a sonic hedgehog inhibitor, recapitulate *in vitro* the major milestones of cortical development, leading to the sequential generation of a diverse repertoire of neurons that display most salient features of genuine cortical pyramidal neurons. When grafted into the cerebral cortex, these neurons develop patterns of axonal projections corresponding to a wide range of cortical layers, but also to highly specific cortical areas, in particular visual and limbic areas, thereby demonstrating that the identity of a cortical area can be specified without any influence from the brain. The discovery of intrinsic corticogenesis sheds new light on the mechanisms of neuronal specification, and opens new avenues for the modelling and treatment of brain diseases.

The cerebral cortex is the most complex structure in the mammalian brain, and displays an unparalleled diversity of neuronal subtypes. This complexity is established sequentially, starting with the emergence of the cortical primordium in the forebrain. Forebrain identity is thought to constitute a primitive pattern of neural identity, which is retained through the inhibition of caudalizing morphogen signals¹. The forebrain then undergoes patterning along the dorso–ventral axis, mainly through the induction of ventral identity by sonic hedgehog (SHH), secreted from the ventral neural tube and underlying tissue². This leads to the specification of the two main populations of cortical neurons, pyramidal neurons and interneurons, which are generated from the dorsal and the ventral part of the telencephalon, respectively^{1,3–5}.

A next level of complexity emerges through the specification of different subtypes of cortical neurons that will populate specific cortical layers, where they exhibit specific patterns of gene expression and connectivity. This specification follows a coordinated temporal pattern: neurons from different layers are generated sequentially⁶, but the underlying mechanisms remain poorly known^{7–10}. Finally, neurons from distinct cortical areas develop selective patterns of gene expression and projections, through the interplay between factors intrinsic to the cortex and extrinsic factors from the brain and body^{3,11–13}.

A better understanding of cortical development would greatly benefit from reductionist approaches. Here we tested whether the complex developmental processes leading to the formation of the cerebral cortex could be studied in a meaningful way using an embryonic stem cell (ESC)-based system. This led to the surprising discovery of an intrinsic pathway that recapitulates *in vitro* most major milestones of corticogenesis.

Generating cortical progenitors from ESCs

To explore the possibility that forebrain identity might emerge during differentiation of ESCs in default conditions, as during *in vivo* embryogenesis¹, we adapted a method of neurogenesis from adherent monocultures of ESCs¹⁴. We cultured ESCs at low density in a chemically defined default medium (DDM) devoid of serum or

any morphogen but allowing cell survival by insulin. ESCs underwent an efficient neurogenesis process leading first to the generation of nestin⁺ neural progenitors, followed by β -tubulin-III⁺ and MAP2⁺ (also known as MTAP2⁺) neurons, and finally by astroglial production (Fig. 1a and Supplementary Fig. 1). Thus, ESCs cultured in the absence of any added morphogen efficiently differentiate into neural cells, as described^{15,16}, and follow a neuron–glia sequence highly reminiscent of cortical cells *in vitro* and *in vivo*¹⁷.

We next looked at the regional identity of the neural progenitors generated in these default conditions: we found that most neural precursors expressed early markers of the anterior neurectoderm (OTX1 and OTX2, Fig. 1c, n), markers of the dorsal forebrain (OTX1 and PAX6, Fig. 1d, e, n) and to a larger extent markers of the ventral forebrain (GSH2 (also known as GSX2) and NKX2-1, Fig. 1f, g, n and Supplementary Fig. 2a, d). In contrast, only a minority (<1%) of progenitors expressed markers indicative of more caudal identity (Fig. 1n and Supplementary Fig. 2b, c). The forebrain-like identity of ESC-derived neural progenitors was further confirmed by polymerase chain reaction with reverse transcription (RT–PCR) experiments (Fig. 1m).

Because SHH is required for the induction of the ventral identity of the forebrain, we hypothesized that the large proportion of ventral forebrain progenitors in DDM cultures might be induced by an endogenous SHH activity. Consistent with this hypothesis, we found that *Shh* was induced after 4 days of DDM culture (Fig. 1m). We then attempted to block it during the neural progenitor differentiation process by using cyclopamine, a specific antagonist of the SHH pathway¹⁸ (referred to as the ‘DDM plus cyclopamine’ condition). Cyclopamine did not change the neuralization process, nor did it alter the forebrain identity of the neural progenitors (Fig. 1a, b, h, n) or their proliferative pattern (data not shown). However, the addition of cyclopamine caused a massive increase in the expression of dorsal markers, whereas expression of ventral markers was almost abolished (Fig. 1i–n; $P < 0.001$ for all markers). Thus, after inhibition of SHH signalling in DDM cultures, ESCs can be converted to a largely

¹IRIBHM (Institute for Interdisciplinary Research), ²Laboratory of Neurophysiology, Université Libre de Bruxelles (ULB), 808 Route de Lennik, B-1070 Brussels, Belgium. ³Institut de Physiologie et Biologie Cellulaires, Université de Poitiers, Centre National de la Recherche Scientifique (CNRS), 40 avenue du recteur Pineau, Poitiers, F-86022, France.

homogeneous population of neural progenitors expressing markers of the presumptive cerebral cortex (Supplementary Fig. 2e)¹⁹.

Generating pyramidal neurons from ESCs

We next examined the identity of the neurons generated in DDM and DDM plus cyclopamine conditions. We found that most

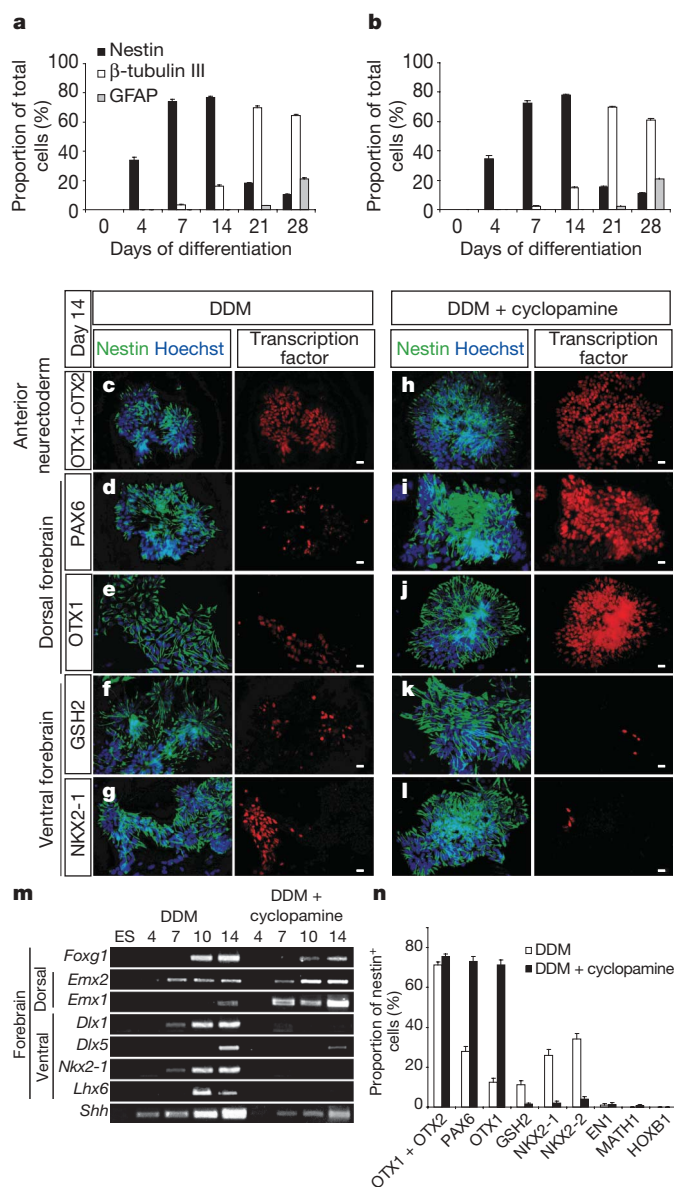


Figure 1 | Differentiation of ESCs into cortical progenitors. **a, b**, Sequential generation of nestin⁺ neural progenitors, β -tubulin-III⁺ neurons and GFAP⁺ astroglial cells, after differentiation in DDM (**a**) or DDM plus cyclopamine (**b**). Histograms show the proportion of cells expressing the given marker among all cells visualized by nuclear staining. No difference is observed between DDM and DDM plus cyclopamine conditions. **c–l**, Expression of the early forebrain regionalization markers OTX1 and OTX2 (**c, e, h, j**), PAX6 (**d, i**), GSH2 (**f, k**) and NKX2-1 (**g, l**) among neural progenitors after 14 days of differentiation in DDM (**c–g**) or DDM plus cyclopamine (**h–l**). Nestin, green; markers, red; nuclear staining, blue. **m**, RT-PCR of specific marker genes of the telencephalon (*Foxg1*), dorsal forebrain (*Emx1* and *Emx2*) and ventral forebrain (*Dlx1*, *Dlx5*, *Nkx2-1* and *Lhx6*) and *Shh* from day 4–14 cultures of cells in DDM or DDM plus cyclopamine. **n**, Proportions of nestin⁺ neural progenitors expressing specific neural regionalization markers after 14 days of differentiation in DDM or DDM plus cyclopamine, showing a conversion from ventral to dorsal forebrain identity in DDM plus cyclopamine. Data are represented as mean \pm s.e.m. Scale bars, 20 μ m.

β -tubulin-III⁺ neurons derived from neural progenitors generated in DDM plus cyclopamine expressed the vesicular glutamate transporters VGLUT1 and VGLUT2 (also known as SLC17A7 and SLC17A6, respectively, Fig. 2d), markers of pyramidal neurons of the cerebral cortex (Fig. 2e), whereas only a minority of them expressed the vesicular GABA (γ -aminobutyric acid) transporter VGAT, a marker of GABAergic inhibitory neurons (Fig. 2f). In the case of neurons generated in DDM, we observed a much greater proportion of GABAergic neurons (Fig. 2a–g; $P < 0.001$, z-test for proportions) and a much lower proportion of glutamatergic neurons ($P < 0.001$ and $P > 0.05$, respectively, for VGLUT1 and VGLUT2). No markers of non-cortical fates, such as tyrosine hydroxylase, choline acetyltransferase, rhodopsin or GABA- α -receptor were found in either DDM or DDM plus cyclopamine conditions (data not shown). We then performed patch-clamp recordings, and found that

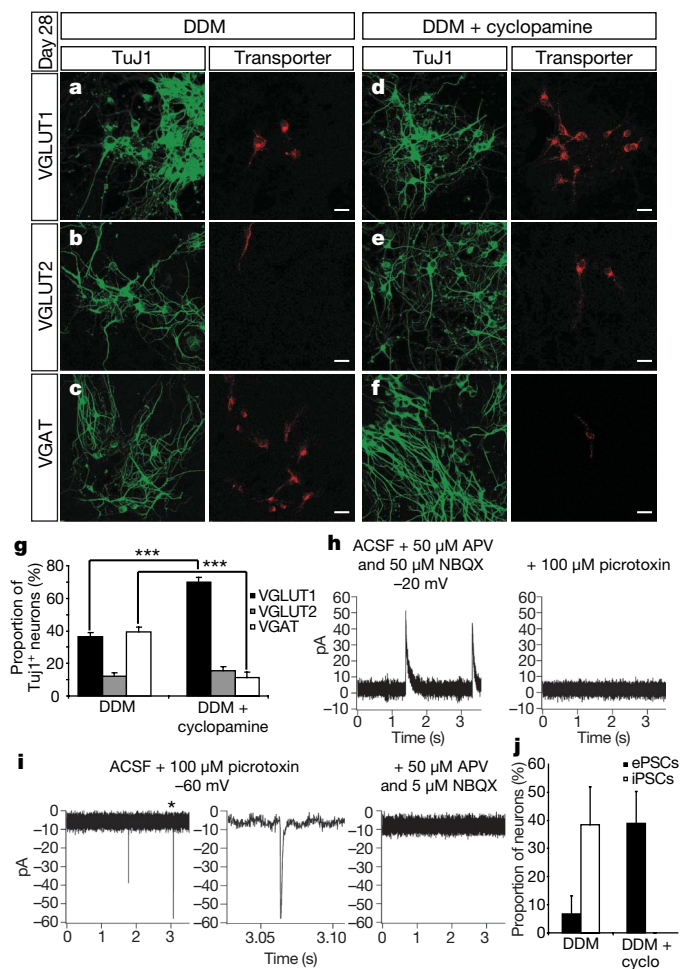


Figure 2 | Generation of functional cortical neurons from ESCs in DDM plus cyclopamine. **a–f**, Expression of glutamatergic vesicular transporters VGLUT1 and VGLUT2 as well as the GABA vesicular transporter VGAT in β -tubulin-III⁺ neurons after 28 days in culture, after differentiation in DDM (**a–c**) or DDM plus cyclopamine (**d–f**). Markers, red; β -tubulin-III (stained with TuJ1), green. Scale bar, 20 μ m. **g**, Proportion of VGLUT1, VGLUT2 and VGAT expression among β -tubulin-III⁺ neurons after 28 days in culture, in DDM or DDM plus cyclopamine conditions. Data are shown as mean \pm s.e.m. ($N = 3$ experiments; $***P < 0.001$, z-test for proportions). **h, i**, Recorded iPSCs (**h**) and ePSCs (**i**) from neurons after 28 days, after differentiation in DDM, showing inhibition by picrotoxin (GABA_A antagonist) and APV/NBQX (NMDA and AMPA receptors antagonists), respectively. The second EPSC in **i** is the EPSC marked with an asterisk in the first graph shown at a different time scale. **j**, Proportion of recorded neurons displaying iPSCs and ePSCs in each condition. Data are shown as mean \pm s.e.m. ($N = 33$ neurons; $P < 0.05$, Fischer exact t-test). ACSF: artificial cerebrospinal fluid.

neurons showed spontaneous action potentials and synaptic currents in both DDM and DDM plus cyclopamine conditions (Fig. 2h, i and Supplementary Fig. 3a–i). However, in DDM conditions, most recorded neurons displayed spontaneous GABAergic inhibitory post-synaptic currents (iPSCs), whereas, in DDM plus cyclopamine conditions, most recorded neurons displayed glutamatergic excitatory post-synaptic currents (ePSCs; Fig. 2j; $P < 0.05$; Fisher exact t -test). Thus, the differences in expression of neurotransmission markers between neurons derived from DDM and those from DDM plus cyclopamine conditions correlate with functional differences in synaptic activity.

Cortical neurons can be also distinguished by their morphological properties: cortical excitatory neurons exhibit a pyramidal or unipolar morphology, characterized by the presence of one dendrite that is wider than the others, whereas GABAergic interneurons show multipolar morphologies. We compared the morphology of the neurons generated in each condition, and observed an important qualitative shift, with many more neurons in DDM plus cyclopamine displaying a unipolar pyramidal morphology (Fig. 3a–c, $P < 0.01$, χ^2 test). We used a pyramidal morphology index (PMI)²⁰ to objectively distinguish the dendritic morphology of cortical pyramidal neurons and inhibitory neurons. Most neurons generated in DDM plus cyclopamine displayed a PMI corresponding to pyramidal neurons that was significantly higher than that in DDM conditions (Fig. 3d; $P < 0.001$, z -test for proportions; Supplementary Fig. 4; $P < 0.001$; Mann–Whitney test).

Within the cortex, glutamatergic pyramidal neurons are polarized and oriented radially, with one large dendrite pointing towards the cortical surface and the axon pointing towards the ventricular surface. Using a slice overlay assay²¹ we tested whether ESC-derived neurons can orient radially like pyramidal neurons. Neurons were first differentiated *in vitro* from green fluorescent protein (GFP)⁺ ESCs, and then dissociated and plated on mouse postnatal cortical slices. Most neurons generated in DDM plus cyclopamine displayed a radial orientation with a dendrite pointing to the cortical surface (Fig. 3e–h), whereas only a minor fraction of neurons generated in DDM showed this behaviour (Fig. 3h; $P < 0.001$).

Together, these data indicate that most of the neurons generated from ESCs in DDM plus cyclopamine display morphological and functional properties characterizing pyramidal neurons, whereas neural progenitors produced in DDM produces a mixture of GABAergic multipolar neurons and pyramidal neurons, in marked correlation with the differential dorso–ventral identity of the neural progenitors generated in each condition.

Sequential cortical neurogenesis from ESCs

Pyramidal neurons can be subdivided into numerous subtypes, defined *in vivo* by a repertoire of molecular markers, identifying distinct layer-specific neuronal populations^{7,22–24}. To examine the diversity of the neuronal repertoire derived from ESCs in DDM plus cyclopamine, we studied the expression of a large set of layer-specific markers corresponding to the major subtypes of cortical neurons generated *in vivo* in the mouse (Fig. 4a and Supplementary Fig. 5). We determined the expression of these markers, alone or in combination, at the peak of generation of neurons (day 21) and found that each was expressed by a specific subset of the neurons, according to a combinatorial pattern notably similar to the one observed for native cortical neurons (Supplementary Figs 5 and 6a–i and Supplementary Table 2). Although the repertoire of ESC-derived cortical neurons closely matches the one observed *in vivo*, we noted some quantitative differences, in particular an under-representation of upper layer subtypes (Fig. 4b, Supplementary Fig. 5 and Supplementary Table 2). This bias is also observed in cultures of genuine cortical progenitors⁸, suggesting that *in vivo* cues are required for the full generation of upper layer neurons²⁵.

These results indicate that a diverse and physiological repertoire of distinct subtypes of neurons corresponding to deep and upper cortical

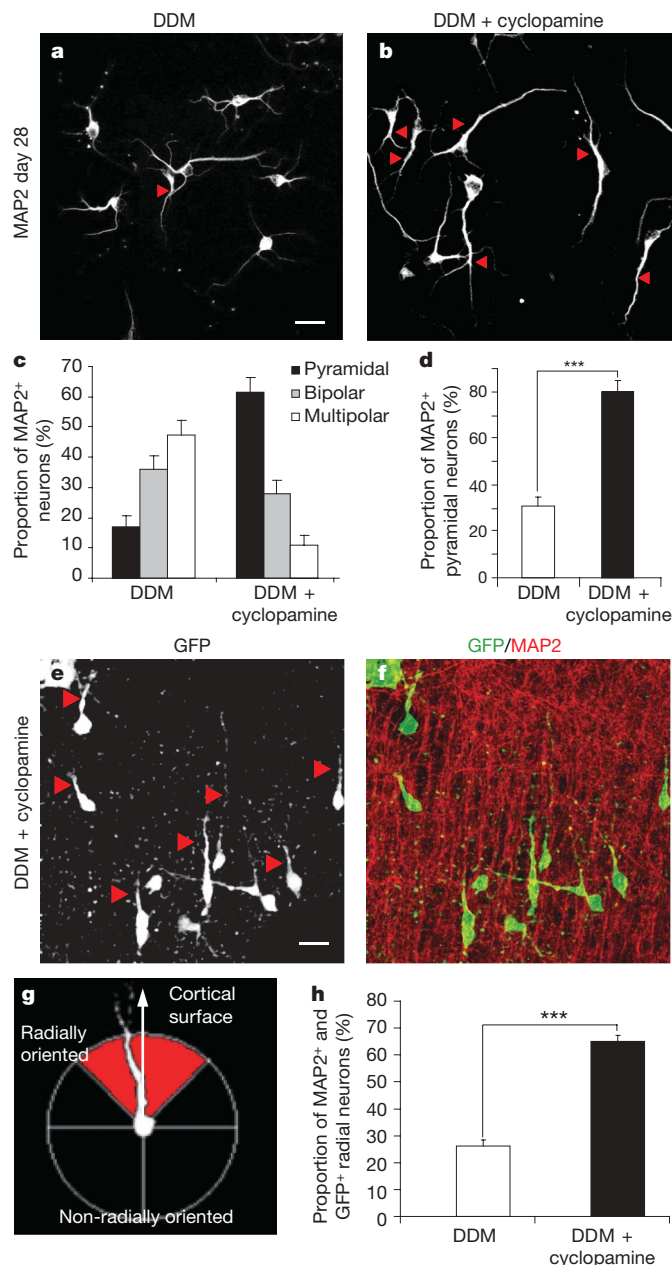


Figure 3 | ESC-derived neurons in DDM plus cyclopamine display morphological features of pyramidal neurons. **a, b**, Typical morphology observed for MAP2⁺ mature neurons cultured for 28 days, after differentiation in DDM (**a**) or DDM plus cyclopamine (**b**). Red arrowheads point to large dendrites that are characteristic of pyramidal neurons. Scale bar represents 20 μ m. **c**, Distribution of neuron morphology. Semi-quantitative scoring of neurons displaying pyramidal, unipolar or multipolar morphology ($N = 3$ experiments; $P < 0.001$ χ^2 test). **d**, Proportion of neurons displaying a PMI above the cut off of 1.2 after DDM or DDM plus cyclopamine conditions, showing a conversion to a pyramidal morphology in DDM plus cyclopamine ($N = 3$ experiments; $***P < 0.001$, z -test for proportions). Orientation pattern of neurons derived from GFP⁺ ESCs in DDM plus cyclopamine (**e–h**) conditions, differentiated for 21 days, and cultured for 4 days on top of postnatal cortical slices. GFP is in white (**e**) and green (**f**) and MAP2 in red in **f**. Apical surface of the cortical slices is up. Dendrites that are located apically and oriented radially are marked by red arrowheads in **e, g, h**. Quantification of the proportion of neurons displaying radial orientation for neurons derived in DDM or DDM plus cyclopamine conditions. Data are represented as mean \pm s.e.m. ($N = 3$ experiments; $***P < 0.001$, z -test for proportions). Scale bar represents 20 μ m. All data are shown as mean \pm s.e.m.

layers are generated from ESCs in DDM plus cyclopamine conditions, raising the question of whether they are specified according to a precise temporal pattern, as observed with native cortical neurons^{8,22} (Fig. 4a). We addressed this issue by examining the timing of onset of expression of these markers throughout the *in vitro* neurogenesis process (Fig. 4b). The first detectable marker to appear in neurons was reelin, after 6 days of differentiation (Fig. 4b). The early appearance of reelin⁺ neurons suggested that these could correspond to Cajal–Retzius neurons (the earliest generated cortical neurons²²), which was confirmed by combined stainings for other markers of Cajal–Retzius neurons and overlay assays (Supplementary Fig. 7). TBR1 was the second marker to be detected (after 7 days of differentiation), followed by OTX1 (8 days), CTIP2 (9 days) and finally SATB2 and CUX1 (12 days). Overall, this sequence is highly similar to the

temporal pattern observed for genuine cortical neurons *in vivo* and *ex vivo* (Supplementary Fig. 5)^{8,22}. Next we performed 5-bromodeoxyuridine (BrdU) pulse–chase experiments combined with staining of the neuronal markers to determine the date of birth of each neuronal subtype in the ESC system (Fig. 4c). This analysis revealed that neurons expressing layer-specific markers are generated in sequential overlapping waves: first reelin and TBR1⁺ neurons (peak at day 10–11), followed by CTIP2⁺ neurons (peaking at day 12–13), and then by CUX1⁺ and SATB2⁺ neurons (peaking at day 14–16; Fig. 4c).

These data demonstrate that ESC-derived neurons corresponding to distinct cortical subtypes appear through a coordinated sequence, as observed *in vivo*^{6,26}. To explore the cellular mechanism involved, we undertook cell lineage experiments using clonal cell analyses⁸. We first performed clonal dilutions of ESC-derived neural progenitors at a

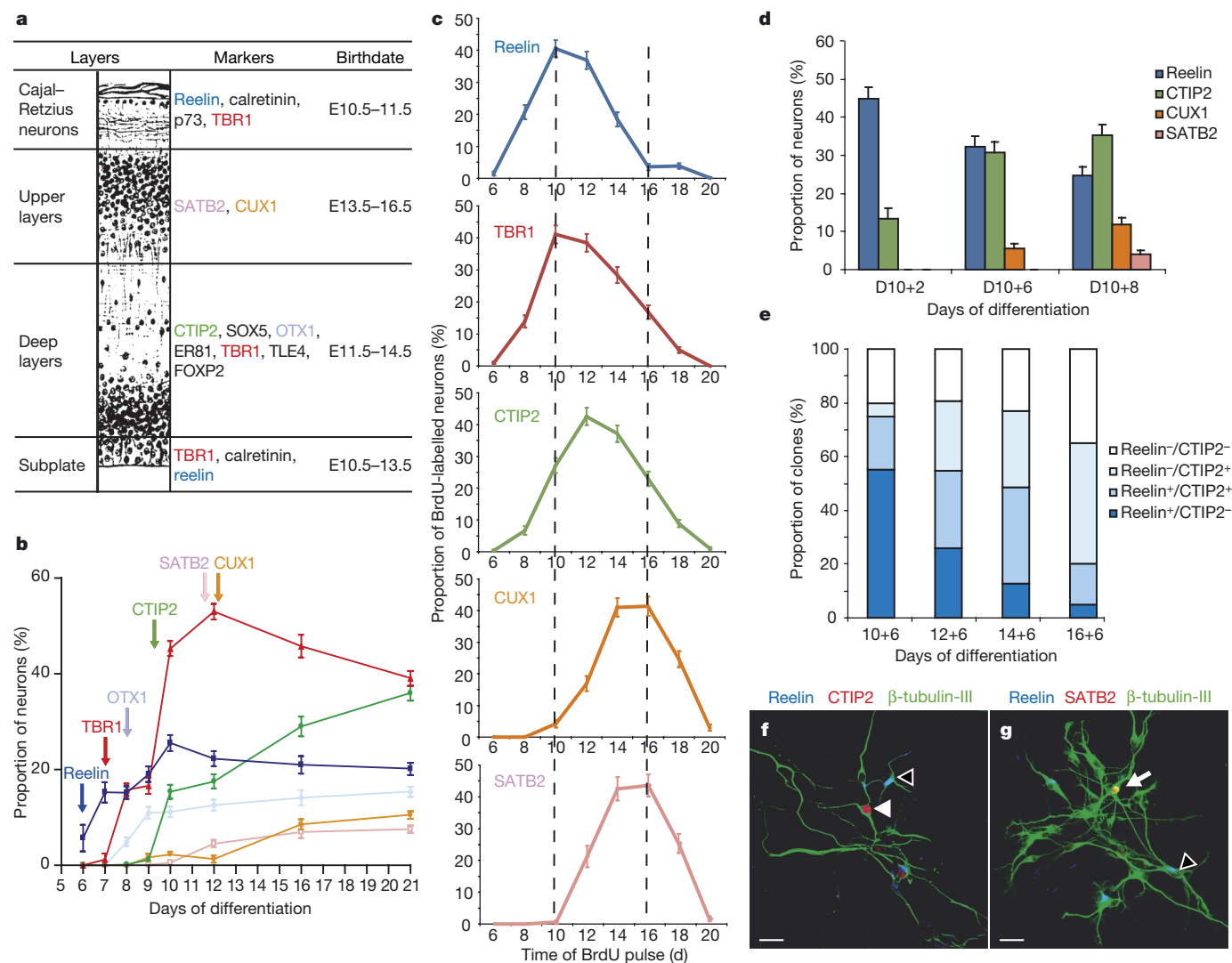


Figure 4 | The sequential generation of the different subtypes of ESC-derived cortical neurons is similar to the *in vivo* situation, and is encoded within single cell lineages. **a**, Scheme depicting patterns of the layer-specific markers in cortical neurons, and their timing of generation *in vivo* (adapted from refs 22, 45). **b**, Evolution in time of the proportion of β-tubulin-III⁺ neurons expressing layer-specific markers in DDM plus cyclopamine. Coloured arrows indicate the first day of appearance of each marker ($N = 3$ experiments). **c**, Birthdating analysis. Cultures were pulse-labelled with BrdU for 24 h at various time points (x axis). Cultures were stopped at day 21, and the proportion of BrdU fully labelled nuclei was quantified among β-tubulin-III⁺ neurons expressing each specific marker ($N = 3$ experiments). **d**, Clonal analysis: ESC-derived neural progenitors were dissociated after 10 days of differentiation, followed by culture at clonal density for 2–8 days. The proportion of β-tubulin-III⁺ neurons expressing

layer-specific markers was then quantified ($N = 3$ experiments). **e**, Clonal analysis reveals the shift in competence of individual neural progenitors. Neural progenitors from ESCs were dissociated at distinct time points (days 10–16 of differentiation), and cultured for 6 days at clonal density, revealing a shift in competence from early progenitors generating mainly reelin⁺ neurons, to later progenitors generating mainly CTIP2⁺ neurons. Values represent the mean number of clones containing at least one marked neuron. **f, g**, Typical example of neuronal clones (stained with β-tubulin-III in green) resulting from dissociation at day 14 followed by 6 days of clonal culture, containing reelin (blue, open arrowhead) and CTIP2 (red, filled arrowhead) neurons (**f**), and reelin (blue, open arrowhead) and SATB2 (red, arrow) neurons (**g**). Scale bar represents 20 μm. All data are represented as mean + s.e.m.

fixed starting point (10 days of differentiation) and characterized the neuronal progeny after serial time points (2, 6 and 8 days, Fig. 4d). At early stages (2 days of clonal culture), such clones of neural progenitors generated mostly *reelin*⁺ neurons and a few *CTIP2*⁺ neurons. With time, the proportion of *reelin*⁺ neurons decreased substantially, whereas the proportion of *CTIP2*⁺ neurons increased. *SATB2*⁺ and *CUX1*⁺ neurons appeared later in the clones, with a maximum reached after 8 days of clonal culture days *in vitro*. Next we dissociated neural cultures from ESCs at different time points (from day 10 to day 16) and looked at their progeny after 6 days *in vitro* (Fig. 4e). Consistently with the previous data, we observed that clones obtained from early progenitors (dissociated at day 10) generated mostly *reelin*⁺ and some *CTIP2*⁺ neurons after 6 days, whereas clones obtained from later progenitors (dissociated at days 12, 14 and 16) contained fewer and fewer *reelin*⁺ neurons but many more *CTIP2*⁺ neurons. These experiments also enabled detection of a substantial proportion of clones containing different types of neurons, such as *reelin* and *CTIP2* single-labelled neurons (with a maximum at day 14 + 6) and *reelin* and *SATB2* single-labelled neurons (Fig. 4e–g). Together, these results indicate that the ESC intrinsic pathway leads to the generation of multipotent cortical progenitors that change in competence over time, as suggested for genuine cortical stem cells^{8,27}.

Cortical axonal projections of ESC neurons

To characterize further the functionality and identity of ESC-derived neurons, we performed *in vivo* grafting experiments in neonatal mice^{13,28} using a TAU (MAPT)–GFP knock-in ESC line²⁹ to identify neurons and their axonal projections. ESCs were allowed to differentiate for 12–17 days in DDM plus cyclopamine conditions, and then were dissociated and grafted into the frontal cortex of mouse neonates. The differentiation and patterns of projections of the grafted cells were analysed after one month (Fig. 5 and Supplementary Fig. 8). In all cases, the graft was localized underneath the frontal cortex or in the grey matter itself, and most of the grafted cells consisted of GFP⁺ neurons that displayed molecular and morphological landmarks of mature pyramidal neurons, with no evidence of teratoma formation (Fig. 5a–d, Supplementary Figs 9 and 10, Supplementary Note 1 and Supplementary Table 3).

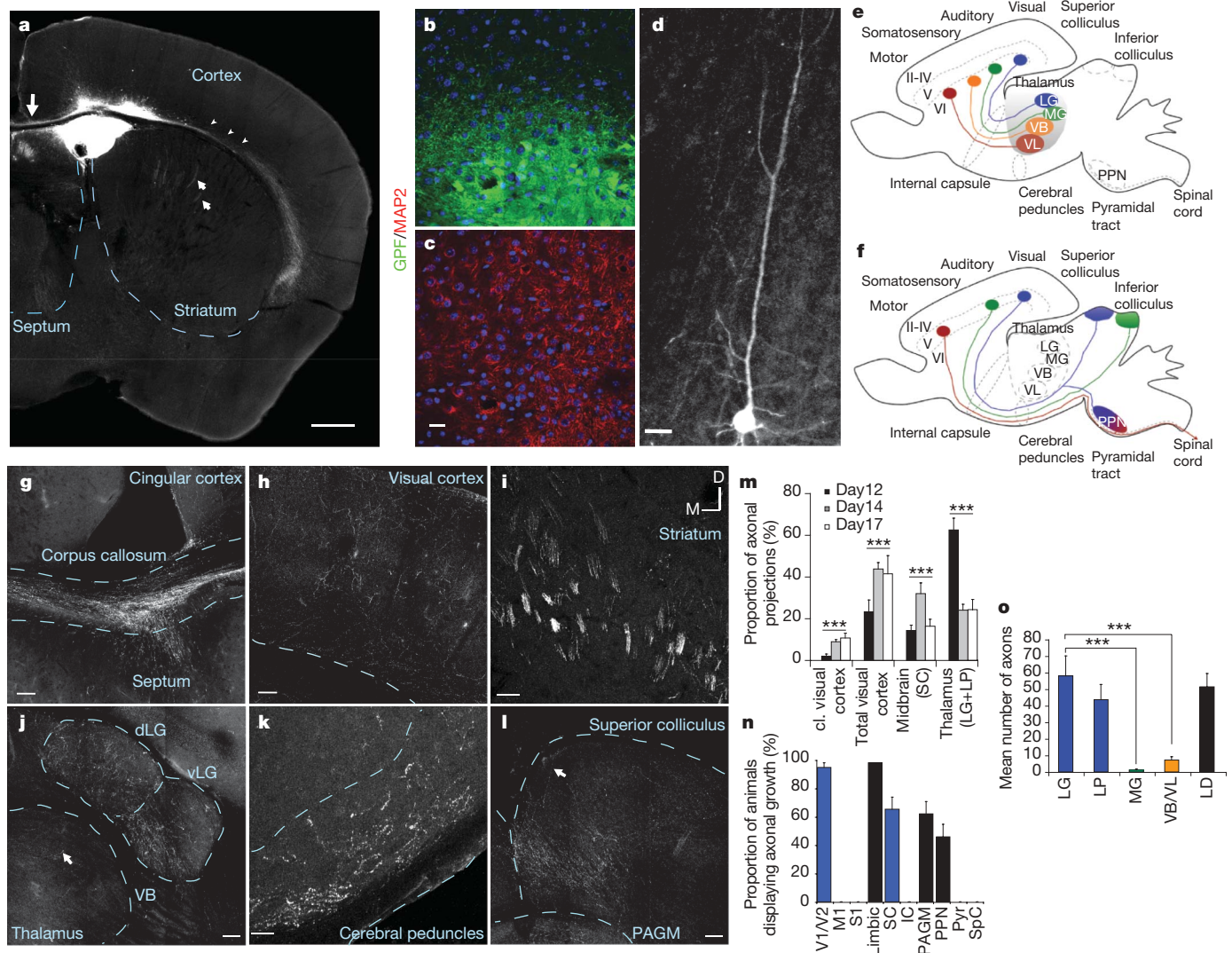
We next examined the GFP⁺ axonal projections of the grafted neurons, and compared them with the endogenous pattern of cortical projections (schematized in Fig. 5e, f). In most of the grafted animals, we found significant projections throughout the external capsule and corpus callosum, up to the ipsilateral and contralateral cortex, corresponding to cortico–cortical projections (Fig. 5g, h, Supplementary Fig. 8a–c and Supplementary Table 4). GFP⁺ projections corresponding to subcortical efferents were found in the striatum, internal capsule, thalamus, cerebral peduncles and midbrain, and to pontine nuclei in the hindbrain (Fig. 5i–l, Supplementary Fig. 8d, e and Supplementary Table 4). Similar patterns of differentiation and axonal projections were observed when ESC-derived neuronal preparations were grafted *in utero* in embryonic day 13 embryos (Supplementary Note 2 and Supplementary Fig. 11d–g).

Overall, the projections of the grafted neurons markedly resemble the pattern of cortical efferents, providing further evidence that most neurons generated from ESCs in DDM plus cyclopamine have a cortical identity. These results also demonstrate that the pattern of projections of the neurons is generated by a diverse repertoire of cortical neurons, including projections to the thalamus (corresponding to cortical layer VI), the midbrain and hindbrain (layer V), and the cerebral cortex (layers II/III and V). Importantly, when performing grafting experiments using neurons differentiated in DDM conditions (thus containing far fewer pyramidal neurons, Figs 1–3), the grafted neurons displayed fewer cortical-like projections, but instead showed local neurite outgrowth and scattered multipolar neurons (Supplementary Fig. 12 and Supplementary Table 6), confirming that the pattern of projections of neurons derived in DDM plus cyclopamine is mainly the result of an *in vitro* intrinsic program.

We next examined whether the grafting of ESC-derived neural progenitors and neurons differentiated after different periods of time, and thus possibly enriched for specific laminar fates, could generate different layer-specific patterns of projections *in vivo*. This analysis revealed that the proportion of projections found in thalamic, midbrain and cortical targets was different depending on the timing of differentiation. Grafts emanating from cells differentiated for a shorter period (12 days) displayed a higher proportion of projections to the thalamus than grafts emanating from cells differentiated for a longer period (days 14–17). Conversely, grafts emanating from later time points (14–17 days) displayed a higher proportion of projections to the cortex than the grafts from earlier time points (12 days), whereas the proportion of projections to the midbrain followed an intermediate temporal pattern (Fig. 5m; $P < 0.001$, χ^2 test). In correlation with the increase in the proportion of cortico–cortical projections for cells cultured for longer periods, the proportion of *CUX1*⁺ neurons, a marker of cortico–cortical projection neurons, was higher in the grafts emanating from cells cultured for 14 days than those cultured for 12 days *in vitro* ($3 \pm 1\%$ versus $11 \pm 2\%$ (mean \pm s.e.m.), $P < 0.01$). Thus, the intrinsic pathway of corticogenesis leads to the sequential generation of a diverse array of cortical neurons displaying layer-specific patterns of axonal projections.

In addition to layer-specific patterns, cortical neurons display area-specific projections, so that each cortical area projects to selective targets in the rest of the brain³⁰ (Fig. 5e, f). To examine this crucial aspect of cortical identity, we performed a detailed analysis of the patterns of projections of ESC-derived grafted neurons. We found that grafted neurons projecting to the cortex, thalamus and midbrain/hindbrain did not innervate these structures in a diffuse way, but in a precise manner, corresponding mainly to visual and limbic occipital cortex (Fig. 5, Supplementary Fig. 8 and Supplementary Table 4). In the cortex, GFP⁺ axons could be seen in most grafted animals in the visual areas (V1/V2), as well as in the limbic cortex, but no projections to motor (M1) or somatosensory (S1) areas were observed in any grafted animal (Fig. 5n and Supplementary Table 4). In the thalamus (Fig. 5j, o and Supplementary Table 4), GFP⁺ axons were observed to innervate the visual thalamic nuclei in most animals, whereas they were observed to innervate the other primary nuclei of the thalamus in only a few cases. Quantification of the distribution of GFP⁺ fibres in each main thalamic nucleus (Fig. 5o) revealed a much denser innervation in visual thalamic nuclei than in any other primary thalamic nucleus. Similarly, when examining projections to midbrain and hindbrain, in most cases GFP⁺ axons projected to the superior colliculus (Fig. 5l, n), the midbrain target of the visual cortex, whereas no fibres could be observed in the pyramidal tract or spinal cord corresponding to motor projections, or to the auditory inferior colliculus. In addition, most grafted animals sent projections to the limbic cortex, medial thalamic nuclei and midbrain periventricular grey matter, corresponding to structures receiving diffuse projections from the neocortex and limbic cortex (Fig. 5n, Supplementary Fig. 8 and Supplementary Table 4). A similar area-specific pattern of projection was found after different days of *in vitro* differentiation (days 12–17), suggesting that cells were committed to these specific areal fates after at least 12 days *in vitro*.

The projections of ESC-derived neurons look strikingly similar to those described with late embryonic occipital cortex grafted into neonatal frontal cortex (which display a mainly visual identity), but are quite distinct from the pattern of embryonic frontal cortex grafted into frontal cortex (which develop mainly motor-like projections)^{28,31,32}. Importantly, our results were all obtained with grafts in the frontal cortex, suggesting that the patterns observed were not due to the re-specification of the grafted neurons through the influence of the host^{13,31,33}. ESC-derived neural cells thus appear to undergo *in vitro* an area-specific differentiation process that results in a surprisingly specific identity, corresponding mainly to occipital/visual cortex. Confirming this hypothesis, we found that even before grafting most



ESC-derived neural progenitors expressed the CoupTFI (NR2F1) and CoupTFII (NR2F2) transcription factors ($81.64 \pm 2.22\%$ and $95.1 \pm 1.23\%$, respectively, Supplementary Fig. 8f, g), which are preferentially expressed in the embryonic occipital/visual cortex^{34,35}.

Discussion

ESCs constitute a powerful resource in neurobiology, enabling the directed specification of spinal, hindbrain and midbrain neurons^{36–38}. The ability for mouse ESCs to generate forebrain-like progenitors was reported previously, but the identity of the neurons generated

the superior colliculus and PPN. **g**–**l**, Patterns of projections of the grafted ESC-derived neurons in the corpus callosum (**g**), the visual cortex (**h**), the internal capsule, the striatum (**i**) and the thalamus (**j**), with many fibres in LG but very few in the VB (arrow in **j**), the cerebral peduncles (**k**) and the superior colliculus (arrow, **l**). PAGM, Peri-aqueductal gray matter.

m, Quantification of the proportion of fibres found in layer-specific targets, from grafted cells differentiated for different time periods *in vitro*. Increasing the time of differentiation yields more cortico-cortical fibres but less cortico-thalamic fibres (***) $P < 0.001$. **n**, Quantification of the proportion of grafted animals displaying GFP⁺ fibres in cortical areas and subcortical targets. IC, inferior colliculus; SpC, spinal cord; SC, superior colliculus. M1, motor area; Pyr, pyramidal tract; S1, somatosensory area; V1, V2, visual areas. **o**, Quantification of the number of fibres found in selective thalamic nuclei. Area-specific colour codes are the same as in **e** and **f**, and diffuse/limbic targets are labelled in black (***) $P < 0.001$. All values are displayed as means \pm s.e.m. Scale bars represent 500 μ m (**a**), 50 μ m (**b**, **c**, **i**, **k**), 20 μ m (**d**), or 100 μ m (**g**, **h**, **j**, **l**). Dorsal (D) is up and medial (M) is left.

had remained uncertain^{39–41}. Here we demonstrate that ESCs can be efficiently specified into a repertoire of neurons that display cardinal features of cortical neurons, including patterns of axonal projections, opening new avenues to apply ESC technology to cortical biology and diseases⁴².

Corticogenesis from ESCs follows the major milestones of cortical development, including regional and temporal patterning of pyramidal neuron specification, through the generation of multipotent cortical progenitors that sequentially produce neurons displaying distinct layer-specific identities. In contrast, ESC-derived cortical

progenitors have a much more restricted capacity to develop diverse area-specific identities, as most of the generated neurons display a selective pattern of visual/limbic areas. This surprising finding indicates that a specific cortical area can differentiate without any influence from the brain, through an intrinsic pathway that may correspond to an ancestral developmental program of the mammalian forebrain^{43,44}.

METHODS SUMMARY

Embryonic stem cells were propagated using standard procedures. For neural differentiation, ESCs were plated at low density on gelatin-coated dishes and cultured in DDM, devoid of any morphogen or serum derivative. Cyclopamine was added from day 2 to day 10 of differentiation. After 10–14 days of differentiation, neural progenitors were dissociated and plated on polylysine/laminin and cultured in N2B27 medium. For clonal analyses, neural progenitors were dissociated and cultured on polylysine/laminin-coated coverslips at clonal density. Immunofluorescent and RT–PCR detection of the markers were performed according to standard methods. Pyramidal morphology analysis and overlay assays were performed as described^{20,21}, and electrophysiology was performed using standard procedures. *In vivo* grafting was performed as described for embryonic cortical neurons³¹ using neonatal or embryonic mice.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 8 June; accepted 30 July 2008.
Published online 17 August 2008.

- Wilson, S. W. & Houart, C. Early steps in the development of the forebrain. *Dev. Cell* **6**, 167–181 (2004).
- Gunhaga, L., Jessell, T. M. & Edlund, T. Sonic hedgehog signaling at gastrula stages specifies ventral telencephalic cells in the chick embryo. *Development* **127**, 3283–3293 (2000).
- Sur, M. & Rubenstein, J. L. Patterning and plasticity of the cerebral cortex. *Science* **310**, 805–810 (2005).
- Schuurmans, C. & Guillemot, F. Molecular mechanisms underlying cell fate specification in the developing telencephalon. *Curr. Opin. Neurobiol.* **12**, 26–34 (2002).
- Gotz, M. & Sommer, L. Cortical development: the art of generating cell diversity. *Development* **132**, 3327–3332 (2005).
- Bayer, S. & Altman, J. *Neocortical Development* (Raven Press, 1991).
- Molyneux, B. J., Arlotta, P., Menezes, J. R. & Macklis, J. D. Neuronal subtype specification in the cerebral cortex. *Nature Rev. Neurosci.* **8**, 427–437 (2007).
- Shen, Q. *et al.* The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nature Neurosci.* **9**, 743–751 (2006).
- McConnell, S. K. & Kazanowski, C. E. Cell cycle dependence of laminar determination in developing neocortex. *Science* **254**, 282–285 (1991).
- Frantz, G. D. & McConnell, S. K. Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron* **17**, 55–61 (1996).
- Vanderhaeghen, P. & Polleux, F. Developmental mechanisms patterning thalamocortical projections: intrinsic, extrinsic and in between. *Trends Neurosci.* **27**, 384–391 (2004).
- Rash, B. G. & Grove, E. A. Area and layer patterning in the developing cerebral cortex. *Curr. Opin. Neurobiol.* **16**, 25–34 (2006).
- O'Leary, D. D., Chou, S. J. & Sahara, S. Area patterning of the mammalian cortex. *Neuron* **56**, 252–269 (2007).
- Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. & Smith, A. Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nature Biotechnol.* **21**, 183–186 (2003).
- Smukler, S. R., Runciman, S. B., Xu, S., & van der Kooy, D. Embryonic stem cells assume a primitive neural stem cell fate in the absence of extrinsic influences. *J. Cell Biol.* **172**, 79–90 (2006).
- Munoz-Sanjuan, I. & Brivanlou, A. H. Neural induction, the default model and embryonic stem cells. *Nature Rev. Neurosci.* **3**, 271–280 (2002).
- Qian, X. *et al.* Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron* **28**, 69–80 (2000).
- Chen, J. K., Taipale, J., Cooper, M. K. & Beachy, P. A. Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes Dev.* **16**, 2743–2748 (2002).
- Inoue, T., Nakamura, S. & Osumi, N. Fate mapping of the mouse prosencephalic neural plate. *Dev. Biol.* **219**, 373–383 (2000).
- Hand, R. *et al.* Phosphorylation of Neurogenin2 specifies the migration properties and the dendritic morphology of pyramidal neurons in the neocortex. *Neuron* **48**, 45–62 (2005).
- Polleux, F., Morrow, T. & Ghosh, A. Semaphorin 3A is a chemoattractant for cortical apical dendrites. *Nature* **404**, 567–573 (2000).
- Hevner, R. F. *et al.* Beyond laminar fate: toward a molecular classification of cortical projection/pyramidal neurons. *Dev. Neurosci.* **25**, 139–151 (2003).
- Britanova, O. *et al.* Satb2 is a postmitotic determinant for upper-layer neuron specification in the neocortex. *Neuron* **57**, 378–392 (2008).
- Alcamo, E. A. *et al.* Satb2 regulates callosal projection neuron identity in the developing cerebral cortex. *Neuron* **57**, 364–377 (2008).
- Dehay, C. & Kennedy, H. Cell-cycle control and cortical development. *Nature Rev. Neurosci.* **8**, 438–450 (2007).
- Hevner, R. F. From radial glia to pyramidal-projection neuron: transcription factor cascades in cerebral cortex development. *Mol. Neurobiol.* **33**, 33–50 (2006).
- Noctor, S. C., Martinez-Cerdeno, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nature Neurosci.* **7**, 136–144 (2004).
- Gaillard, A., Gaillard, F. & Roger, M. Neocortical grafting to newborn and adult rats: developmental, anatomical and functional aspects. *Adv. Anat. Embryol. Cell Biol.* **148**, 1–86 (1998).
- Wernig, M. *et al.* Tau EGFP embryonic stem cells: an efficient tool for neuronal lineage selection and transplantation. *J. Neurosci. Res.* **69**, 918–924 (2002).
- Paxinos, G. *The Rat Nervous System* (Academic, 1995).
- Pinaudeau, C., Gaillard, A. & Roger, M. Stage of specification of the spinal cord and tectal projections from cortical grafts. *Eur. J. Neurosci.* **12**, 2486–2496 (2000).
- Ebrahimi-Gaillard, A., Guitet, J., Garnier, C. & Roger, M. Topographic distribution of efferent fibers originating from homotopic or heterotopic transplants: heterotopically transplanted neurons retain some of the developmental characteristics corresponding to their site of origin. *Brain Res. Dev. Brain Res.* **77**, 271–283 (1994).
- Barbe, M. F. & Levitt, P. Age-dependent specification of the corticocortical connections of cerebral grafts. *J. Neurosci.* **15**, 1819–1834 (1995).
- Armentano, M. *et al.* COUP-TFI regulates the balance of cortical patterning between frontal/motor and sensory areas. *Nature Neurosci.* **10**, 1277–1286 (2007).
- Sansom, S. N. *et al.* Genomic characterisation of a Fgf-regulated gradient-based neocortical protomap. *Development* **132**, 3947–3961 (2005).
- Lee, S. H., Lumelsky, N., Studer, L., Auerbach, J. M. & McKay, R. D. Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nature Biotechnol.* **18**, 675–679 (2000).
- Wichterle, H., Lieberam, I., Porter, J. A. & Jessell, T. M. Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**, 385–397 (2002).
- Andersson, E. *et al.* Identification of intrinsic determinants of midbrain dopamine neurons. *Cell* **124**, 393–405 (2006).
- Bibel, M. *et al.* Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nature Neurosci.* **7**, 1003–1009 (2004).
- Glaser, T. & Brustle, O. Retinoic acid induction of ES-cell-derived neurons: the radial glia connection. *Trends Neurosci.* **28**, 397–400 (2005).
- Watanabe, K. *et al.* Directed differentiation of telencephalic precursors from embryonic stem cells. *Nature Neurosci.* **8**, 288–296 (2005).
- Gaillard, A. *et al.* Reestablishment of damaged adult motor pathways by grafted embryonic cortical neurons. *Nature Neurosci.* **10**, 1294–1299 (2007).
- Puelles, L. Thoughts on the development, structure and evolution of the mammalian and avian telencephalic pallium. *Phil. Trans. R. Soc. Lond. B* **356**, 1583–1598 (2001).
- Frost, S. B., Milliken, G. W., Plautz, E. J., Masterton, R. B. & Nudo, R. J. Somatosensory and motor representations in cerebral cortex of a primitive mammal (*Monodelphis domestica*): a window into the early evolution of sensorimotor cortex. *J. Comp. Neurol.* **421**, 29–51 (2000).
- von Economo, C. & Koskinas, G. *The Cytoarchitectonics of the Adult Human Cortex* (Springer, 1925).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank G. Vassart, M. Pandolfo and members of the laboratory and IRIBHM for support and discussions. We are indebted to A. Bilheu for technical assistance, J.-M. Vanderwinden for help with confocal microscopy, and V. De Maertelaer for statistical analyses. We thank F. Polleux, B. Hassan and C. Blanpain for comments on the manuscript. We are grateful to S. Arber, A. Goffinet, R. Hevner, R. di Lauro, Y. Sasai, S. Stifani, M. Studer and V. Tarabykin for providing us with antibodies, and to Y.-A. Barde for providing Tau–GFP ESC lines. This work was funded by the Belgian FNRS, the Action de Recherches Concertées (ARC) Programs (to P.V. and S.N.S.), the Interuniversity Attraction Poles Program (IUAP), Belgian State, Federal Office, the Walloon Region Excellence Program CIBLES, the Belgian Queen Elizabeth Medical Foundation and a UCB Neuroscience Award (to P.V.), the Tournesol FNRS/CNRS Program (to P.V. and A.G.), Télévie (to S.N.S.), and a Marie Curie Grant (to T.B.). P.V. is a Senior Research Associate of the FNRS, and N.G., R.H., T.B., J.D. and L.P. were funded as Research Fellows of the FNRS.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.V. (pierre.vanderhaeghen@ulb.ac.be).

METHODS

ESC culture. Embryonic stem cells (E14Tg2a (obtained from Baygenomics), E14Tg2a–GFP, a derivative thereof, and the TAU–GFP knock-in line²⁹) were routinely propagated using standard procedures¹⁴ on gelatin-coated cell culture plastic dishes (0.1% gelatin, Sigma; for E14Tg2a cell lines) or on mitotically inactivated mouse embryonic fibroblasts (MEFs; for the TAU–GFP KI cell line).

Neural differentiation. The neural differentiation protocol was adapted from the adherent monoculture protocol described previously¹⁴. Similar results were obtained with several ESC lines, including distinct derivatives of E14 and J1 cells. ESCs were trypsinized, dissociated and plated at a density of 5×10^3 cells per cm^2 on gelatin-coated cell culture plastic dishes in ESC medium. After adhesion, medium was changed to DDM. DDM consists of DMEM/F12 (Invitrogen–Gibco) supplemented with $1 \times \text{N2}$ supplement ($100 \times \text{N2}$ supplement consists of 8.61 μM insulin, 1 mM transferrin, 2 μM progesterone, 10.01 mM putrescine and 3.01 μM selenite; Invitrogen–Gibco), 2 mM glutamine, $1 \times \text{MEM}$ –nonessential amino acids, 1 mM sodium pyruvate, 0.5 mg ml^{-1} bovine serum albumin (BSA) fraction V (all from Invitrogen–Gibco), and 110 μM β -mercaptoethanol (Sigma). Cyclopamine (Calbiochem) was added from day 2 to day 10 in the differentiation medium at a final concentration of 1 μM . After 10 to 14 days of differentiation, cells were trypsinized, dissociated and plated on polylysine/laminin (Becton–Dickinson) coated glass coverslips and allowed to grow for 4–14 days in N2B27 medium. N2B27 medium consists of a 1:1 mixture of DMEM/F12 supplemented with $1 \times \text{N2}$, 2 mM glutamine, 0.5 mg ml^{-1} BSA fraction V and 110 μM β -mercaptoethanol with Neurobasal supplemented with B27 (without vitamin A; Invitrogen–Gibco) and 2 mM glutamine.

Immunofluorescence and quantification. Medium was removed and cells were rinsed with PBS and then fixed for 30 min in 4% paraformaldehyde (PFA) in PBS, pH 7.4, at 4 °C. PFA was removed and cells were rinsed with PBS. Blocking and permeabilization was done in PBS with 5% horse serum (Invitrogen–Gibco), 3% BSA (Sigma) and 0.3% Triton X-100 (Sigma). Incubation with primary antibodies was done in PBS with 1% horse serum, 3% BSA and 0.1% Triton X-100 overnight (12–18 h) at 4 °C. Cells were rinsed three times with PBS and incubated with secondary antibodies in PBS with 1% horse serum, 33% BSA and 0.1% Triton X-100 for one hour at room temperature (18–20 °C), and then rinsed again with PBS. For multiple labellings, the process was repeated. Nuclei were stained with bisbenzimidazole (Hoechst 33258; Sigma). For some antibodies, antigen retrieval was used before blocking. The primary and secondary antibodies used are described in Supplementary Methods 1.

For quantification of the proportion of cells expressing a specific marker, at least 300 cells from three independent experiments (at least 100 cells from each experiment) were counted for each time point studied.

BrdU labelling. For pulse–chase experiments, BrdU (Becton Dickinson) was added to the differentiation medium on the different days (see text) during 24 h at a final concentration of 10 μM . Cells were dissociated after 21 days of

differentiation and plated on polylysine/laminin-coated coverslips for an additional day. For short-term BrdU uptake experiments, BrdU was added to the differentiation medium during 1 h before cell dissociation and replating. BrdU was detected by immunofluorescence as described above.

Clonal analyses. ESC-derived neural progenitors at day 10–16 were trypsinized and thoroughly dissociated with a Pasteur pipette and subsequently with a 200- μl tip into a single-cell suspension. Dissociation was checked under a brightfield microscope and usually yields >99% single cells. Cells were plated on polylysine/laminin-coated coverslips at clonal density (3–5 cells per mm^2) into clonal medium supplemented with fibroblast growth factor (FGF2). Clonal medium was adapted from ref. 8 and consists in Neurobasal supplemented with $1 \times \text{N2}$, $1 \times \text{B27}$, 2 mM glutamine and $1 \times$ penicillin–streptomycin (Gibco). Half of the medium was changed every three days. FGF2 was added to the medium at concentrations (0.1 ng ml^{-1} or 10 ng ml^{-1}) that have been shown to sustain clonal growth from single cortical progenitors and to allow the sequential generation of the different populations of cortical neurons⁸.

RT–PCR. RT–PCR was performed according to Supplementary Methods 2.

Morphology. For the morphology study, we used the PMI described previously²⁰. Pictures of at least 100 neurons stained for MAP2 were taken randomly for each condition and from two different experiments, and the quantification was done blindly in Adobe Photoshop Software using a 25 μm circle drawn on a superposed layer.

Overlay assay. Postnatal day 2 mice brains were dissected in cold L15 buffer supplemented with glucose, embedded in 3% low-melting point agarose in L15, and sectioned (coronal) on a vibratome at 300 μm . Slices were cultured on a transparent porous membrane (1 μm pore size, Greiner) in a 35-mm well containing N2B27 medium. Neurons derived from GFP-expressing ESCs were dissociated after 21 days by trypsinization and mechanical trituration, washed and resuspended at 1×10^6 cells per millilitre before being plated onto cortical slices. Slices were cultured for 4 days after plating, and then fixed with 4% PFA overnight at 4 °C and processed for immunofluorescence. PFA was removed and slices were rinsed three times with PBS. Immunofluorescence was performed as described previously. Quantification of radial dendrite orientation was done blindly according to ref. 21 using Adobe Photoshop software. At least 300 cells on at least 9 different slices coming from 3 independent experiments were assessed (at least 100 cells from each experiment).

Electrophysiology. Electrophysiological recordings were performed according to standard procedures (Supplementary Methods 3).

In vivo grafting experiments. *In vivo* grafting in neonatal (day of birth or one day later) mice of ESC-derived cortical neurons (after 12–17 days in DDM plus cyclopamine or DDM) was performed as described previously for genuine embryonic cortical neurons³¹, with modifications described in Supplementary Methods 4. *In utero* grafting experiments were performed according to Supplementary Methods 5. Analysis and quantification of the grafts and their projections are described in Supplementary Methods 6.

ARTICLES

Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains

Yusuke Sato^{1,2}, Azusa Yoshikawa^{1,2}, Atsushi Yamagata¹, Hisatoshi Mimura¹, Masami Yamashita^{1,4}, Kayoko Ookata², Osamu Nureki^{2,†}, Kazuhiro Iwai⁵, Masayuki Komada³ & Shuya Fukai^{1,4}

Deubiquitinating enzymes (DUBs) remove ubiquitin from conjugated substrates to regulate various cellular processes. The Zn²⁺-dependent DUBs AMSH and AMSH-LP regulate receptor trafficking by specifically cleaving Lys 63-linked polyubiquitin chains from internalized receptors. Here we report the crystal structures of the human AMSH-LP DUB domain alone and in complex with a Lys 63-linked di-ubiquitin at 1.2 Å and 1.6 Å resolutions, respectively. The AMSH-LP DUB domain consists of a Zn²⁺-coordinating catalytic core and two characteristic insertions, Ins-1 and Ins-2. The distal ubiquitin interacts with Ins-1 and the core, whereas the proximal ubiquitin interacts with Ins-2 and the core. The core and Ins-1 form a catalytic groove that accommodates the Lys 63 side chain of the proximal ubiquitin and the isopeptide-linked carboxy-terminal tail of the distal ubiquitin. This is the first reported structure of a DUB in complex with an isopeptide-linked ubiquitin chain, which reveals the mechanism for Lys 63-linkage-specific deubiquitination by AMSH family members.

Ubiquitin is a post-translational modifier that regulates a wide variety of biological processes^{1–3}. Activated ubiquitin is typically attached to substrate lysine residues by an isopeptide linkage between the terminal carboxyl group of ubiquitin and the ε-amino group of the target lysine residue. Additional ubiquitin molecules can be attached to the lysine residues in ubiquitin itself, producing polyubiquitin chains. Although Lys 48-linked ubiquitin chains are the most abundant linkage *in vivo* and constitute the canonical signal for degradation by the proteasome, all seven lysine residues are used to form polyubiquitin chains *in vivo*³. The roles of ubiquitin chains linked by means of residues other than Lys 48 are less thoroughly understood. However, Lys 63-linked ubiquitin chains serve proteasome-independent roles in a variety of intracellular events, such as DNA repair, ribosomal protein synthesis, inflammatory signalling, endocytosis and vesicular trafficking³. For instance, endosomal cargo proteins such as epidermal growth factor receptor are targets of Lys 63-linked polyubiquitination⁴. Lys 63-linked polyubiquitinated receptors on the cell surface are internalized by endocytosis and delivered to sorting endosomes, where the receptors can be routed back to the cell surface or directed to late endosomes/lysosomes to be degraded⁵.

DUBs are isopeptidases that catalyse the removal of ubiquitin. DUBs can be divided into five groups: ubiquitin-specific protease (USP), ubiquitin carboxy-terminal hydrolase (UCH), Otubain protease (OTU), Machado–Joseph disease protease and JAB1/MPN/Mov34 metalloenzyme (JAMM)^{6,7}. All DUBs except JAMMs are cysteine proteases, whereas JAMMs are zinc metalloproteases.

AMSH (associated molecule with the SH3 domain of STAM, also known as STAMBP) family members^{8,9} are JAMM DUBs, which specifically cleave Lys 63-linked polyubiquitin chains^{10–12}. It has been suggested that AMSH proteins facilitate the recycling of receptors by removing Lys 63-linked polyubiquitin chains, resulting in the trafficking of receptors to the cell surface. The amino-terminal half of AMSH proteins interacts with clathrin and ESCRT (the endosomal sorting complex required for transport) proteins on the endosome,

whereas the C-terminal half is engaged in Lys 63-linkage-specific deubiquitination.

Until now, many DUB structures have been reported, some of which are DUBs in complex with a bound ubiquitin monomer^{13–16}. However, no JAMM DUB structure has been reported in isolation or in complex with ubiquitin, and the mechanism of linkage-specific polyubiquitin cleavage has remained unknown for all DUBs owing to the lack of structures of DUBs in complex with isopeptide-linked ubiquitin chains. Furthermore, there are few structural models that illustrate the basis of linkage-specific ubiquitin binding^{17–19}. Here we report crystal structures of the human AMSH-LP (also known as STAMBP1) JAMM DUB domain alone and in complex with a Lys 63-linked di-ubiquitin (K63-Ub₂) at 1.2 Å and 1.6 Å resolutions, respectively. The complex reveals the structural basis for selective cleavage of Lys 63-linked ubiquitin chains by AMSH family members.

Structure of the AMSH-LP DUB domain

The region containing residues 264–436 of human AMSH-LP is sufficient for the Lys 63-linkage-specific DUB activity (Supplementary Fig. 1). This region is referred to as the DUB domain in this study. The crystal structure of the AMSH-LP DUB domain consists of the ‘JAMM core’ and two AMSH-family-specific insertions, Ins-1 (residues 314–339) and Ins-2 (residues 393–415) (Fig. 1a). The JAMM core has been defined in the crystal structures of the other non-DUB JAMM-motif-containing proteins such as yeast and worm Prp8 (refs 20, 21), human MOV34 (ref. 22) and archaeal AfJAMM (refs 23, 24; Supplementary Fig. 2). In the JAMM core of AMSH-LP, Zn²⁺ is coordinated by His 347, His 349, Asp 360 and a water molecule hydrogen bonded to Glu 292, in a manner similar to that in AfJAMM^{23,24} (Fig. 1b, top). Two important and unique features of the AMSH-LP DUB domain are the insertions Ins-1 and Ins-2, which are conserved among AMSH family members (Supplementary Fig. 3). Ins-1 consists of a pair of antiparallel β-sheets (β4 and β5) and the following α-helix (α2). Ins-2 is composed of one α-helix (α4) and the

¹Structural Biology Laboratory, Life Science Division, Synchrotron Radiation Research Organization and Institute of Molecular and Cellular Biosciences, The University of Tokyo, Tokyo 113-0032, Japan. ²Department of Biological Information, and ³Biological Sciences, Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan. ⁴Department of Medical Genome Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Chiba 277-8501, Japan. ⁵Department of Biophysics and Biochemistry, Graduate School of Medicine, Osaka University, Osaka 565-0871, Japan. [†]Present address: Department of Basic Medical Sciences, The Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan.

following long loop harbouring Cys 402, His 408 and His 410, which coordinate the second Zn^{2+} , together with His 362 on $\alpha 3$ in the JAMM core (Fig. 1b, bottom).

Structure of the complex with K63-Ub₂

For the crystallization of the complex with K63-Ub₂, we prepared the DUB domain of the catalytically inactive AMSH-LP E292A mutant (AMSH-LP^{E292A}) and enzymatically synthesized K63-Ub₂ as described in previous studies^{25–27} (with minor modifications, see Methods). The asymmetric unit contains two complexes; these are almost identical, except for residues 319–322 and 401–414 of AMSH-LP, which seem to be affected by crystal packing. Hereafter, we will discuss the complex that is unaffected by the crystal packing.

In the AMSH-LP^{E292A}•K63-Ub₂ complex, K63-Ub₂ adopts an extended conformation. Recognition of the extended conformation is reasonable because previous NMR studies have shown that Lys 63-linked polyubiquitin chains are extended and lack stable intra-chain ubiquitin–ubiquitin interactions in solution^{27,28}. This is in contrast to Lys 48-linked polyubiquitin chains. There are no intermolecular contacts between AMSH-LP and K63-Ub₂ near Lys 63 (replaced by Arg 63 in the present complex) in the distal ubiquitin and near Gly 76 in the proximal ubiquitin, indicating that AMSH-LP has endo-DUB activity for Lys 63-linked polyubiquitin chains (Fig. 1c). The distal ubiquitin binds to Ins-1 and the JAMM core with a buried surface area of 1,324 Å², which is comparable to those of the other DUB–ubiquitin complexes such as USP7–ubiquitin¹³ (1,831 Å²), UCHL3–ubiquitin¹⁵ (1,166 Å²) and OTU1 (also known as OTUB1)–ubiquitin¹⁶ (1,143 Å²). In contrast, the proximal ubiquitin binds to Ins-2 and the JAMM core with a buried surface area of

472 Å². The β -stranded C-terminal tail of the distal ubiquitin and the conjugated Lys 63 of the proximal ubiquitin are accommodated in a 20-Å-long catalytic groove formed by $\alpha 3$ of the JAMM core and $\beta 5$ of Ins-1 in AMSH-LP (Fig. 1c, d). The Zn^{2+} -coordinating active site is covered with the side chains of Asp 321 in Ins-1 and Phe 407 in Ins-2, resulting in the closed active site (Supplementary Fig. 4).

Proximal ubiquitin recognition

The proximal ubiquitin is recognized by the Zn^{2+} -coordinating loop of Ins-2 and the loop connecting $\beta 6$ and $\alpha 3$ in the JAMM core (Figs 1c and 2). These two regions form a concave surface that binds the molecular surface around Lys 63 of the proximal ubiquitin and ensures that only Lys 63-linked ubiquitin can bind in the proper orientation for catalysis. The interface between AMSH-LP and the proximal ubiquitin has two sequence-specific interactions. First, the O γ atom and main-chain NH of Thr 353 in the JAMM core form a hydrogen bond with the N ϵ and O ϵ atoms of Gln 62 in the proximal ubiquitin, respectively (Fig. 2). Second, the O γ atom and main-chain NH of Ser 358 in the JAMM core form a hydrogen bond with the two O ϵ atoms of Glu 64 in the proximal ubiquitin (Fig. 2). These hydrogen bonds probably allow the tripeptide sequence, -Gln 62-Lys 63-Glu 64-, in the proximal ubiquitin to align correctly for Lys 63-linkage-specific deubiquitination. Furthermore, the JAMM core residue Phe 355 and the Ins-2 residue Phe 407 form the binding site for aliphatic portions of Lys 63 in the proximal ubiquitin (Fig. 2). Finally, the side chain of Pro 409 in Ins-2 provides hydrophobic interactions with that of Phe 4 in the proximal ubiquitin.

The functional importance of proximal ubiquitin recognition by AMSH-LP was assessed by kinetic experiments using T353A, S358A, C402S, F355A and F407A mutants (Supplementary Table 1). Compared to the wild-type enzyme, mutations in the proximal ubiquitin-binding residues and the Zn^{2+} -coordinating Cys 402 in Ins-2 decreased k_{cat} values by 10- to 420-fold with similar or lower K_{m} values. This result indicates that the interaction with the proximal ubiquitin is critical for catalysis, but not for binding affinity. It is unlikely that our mutations in the proximal ubiquitin-binding site reduce k_{cat} values by causing misalignment of the catalytic residues. First, because the side chains of Thr 353 and Ser 358 in AMSH-LP are exposed to solvent, replacement of these residues with Ala is unlikely to alter the structure of the catalytic site significantly. Second, our high-resolution structures show a stable catalytic site with low temperature factors despite the presence of the E292A mutation in the catalytic site. Therefore, we propose that the interaction with the

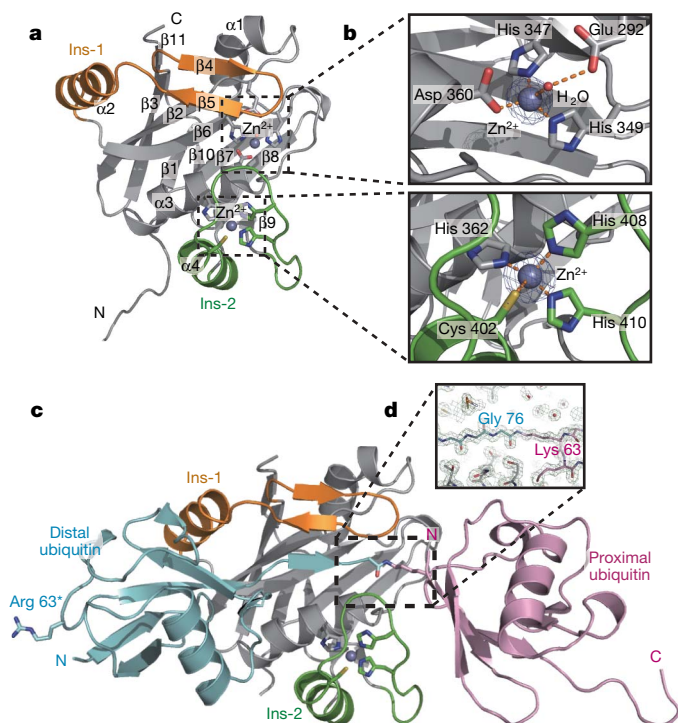


Figure 1 | Overall structure of the AMSH-LP DUB domain alone and in complex with K63-Ub₂. The JAMM core, Ins-1 and Ins-2 of AMSH-LP are coloured grey, orange and green, respectively. The proximal and distal ubiquitins are coloured pink and cyan, respectively. **a**, The AMSH-LP DUB domain. α -helices and β -strands are numbered $\alpha 1$ – $\alpha 4$ and $\beta 1$ – $\beta 11$, respectively. **b**, Anomalous difference Fourier maps in Zn^{2+} -coordinating sites, contoured at 5σ . **c**, The AMSH-LP^{E292A}•K63-Ub₂ complex. Arg 63 in the distal ubiquitin (Lys 63 in the wild type) is shown by a stick model. **d**, A $2F_o - F_c$ electron density map around the isopeptide linkage, contoured at 1.5σ .

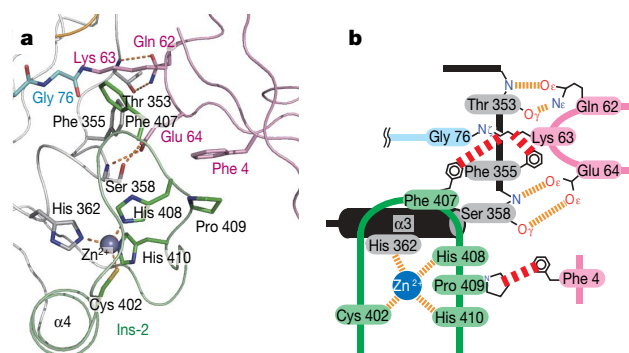


Figure 2 | Proximal ubiquitin recognition. **a**, The interface between AMSH-LP and the proximal ubiquitin. The colouring scheme is the same as in Fig. 1. **b**, Schematic representation of **a**. Hydrogen bonds and hydrophobic interactions are displayed as dashed orange and red lines, respectively. The JAMM core and Ins-2 in AMSH-LP are labelled by grey and green backgrounds, respectively. The proximal and distal ubiquitins are highlighted by pink and cyan backgrounds, respectively.

proximal ubiquitin facilitates the correct orientation of the proximal ubiquitin and accurate positioning of the isopeptide-linked Lys 63 for efficient Lys 63-linkage-specific deubiquitination. Stabilization of the Ins-2 loop by the Zn^{2+} coordination is essential in this context. Lys 48-linked polyubiquitin chains can be cleaved in the presence of the C402S mutant but not of the other mutants (Supplementary Fig. 5).

Distal ubiquitin recognition

The hydrophobic patch surrounding Ile 44 of ubiquitin is of central importance for the ubiquitin binding of numerous ubiquitin-interacting proteins²⁹. Many cysteine-protease DUBs, such as USP7 (ref. 13), UCHL3 (ref. 15) and OTU1 (ref. 16), recognize the hydrophobic patch surrounding Ile 44 in the distal ubiquitin. However, others, including IsoT (also known as USP5)¹⁴, do not. In the present AMSH-LP^{E292A}•K63-Ub₂ complex, the hydrophobic patch of the distal ubiquitin formed by side chains of Leu 8, Ile 44 and Val 70, and the C β atom of His 68 interacts with aliphatic side chains of Val 328 and Phe 332 in Ins-1 of AMSH-LP (Fig. 3a, b). On the face of ubiquitin opposite the Ile 44-centred hydrophobic patch, side chains of Ile 36, Leu 69 and Leu 71, and the C γ atom of Thr 7, form a hydrophobic pocket that accommodates the side chain of Met 370 in the JAMM core of AMSH-LP (Fig. 3a, b). Mutations in these two hydrophobic interfaces, F332A and M370A, increased K_m values 12- and 18-fold, respectively, with little effect on k_{cat} values, in contrast to mutations in the proximal binding site (Supplementary Table 1). This result indicates that the binding affinity primarily relies on the

two hydrophobic interfaces between AMSH-LP and the distal ubiquitin. In addition to hydrophobic interactions, four hydrogen bonds are formed between Ins-1 of AMSH-LP and the distal ubiquitin (Fig. 3a, b). The main-chain NH groups of Val 328 and Thr 342 in AMSH-LP hydrogen bond with the O ϵ atom of Gln 49 and the main-chain CO of Leu 8 in the distal ubiquitin, respectively. The two O ϵ atoms of Glu 329 in AMSH-LP hydrogen bond with the N ζ atom of Lys 48 and the main-chain NH of Gln 49 in the distal ubiquitin. Replacement of Glu 329 by Ala increased the K_m value 2.9-fold (Supplementary Table 1).

AMSH-LP also interacts with the C-terminal tail of the distal ubiquitin (Fig. 3c, d). The C-terminal tail of the distal ubiquitin forms a β -sheet with $\beta 5$ (residues 323–325) in Ins-1 of AMSH-LP. The β -stranded C-terminal tail is further stabilized by an extensive hydrogen bonding network, as shown in Fig. 3d. The main-chain NH and CO of Gly 76 in the distal ubiquitin hydrogen bond with the O δ atom of Asp 360 and the O γ atom of Ser 357 in AMSH-LP, respectively. The main-chain CO groups of Arg 74 and Arg 72 in the distal ubiquitin hydrogen bond with the O γ atom of Thr 363 and the O η atom of Tyr 367 in AMSH-LP, respectively. The N ϵ and N η atoms of Arg 74 in the distal ubiquitin hydrogen bond with the two O δ atoms of Asp 324 in AMSH-LP. In addition to the extensive hydrogen bond network, the side chain of Leu 73 in the distal ubiquitin is accommodated in a hydrophobic pocket formed by the side chains of Cys 294, Ile 296, Met 325 and Trp 345 in AMSH-LP. The side chain of distal ubiquitin Leu 71 is stabilized by hydrophobic interaction with the side chain of Tyr 367. These extensive interactions probably guarantee accurate positioning of the scissile isopeptide bond in the Zn^{2+} -coordinating active site.

Catalytic mechanism

As described above, we mutated the water-mediated Zn^{2+} -coordinating residue Glu 292 to Ala in AMSH-LP to facilitate crystallization of the complex with K63-Ub₂. Thus, we do not observe electron density corresponding to the catalytic Zn^{2+} or to the Zn^{2+} -coordinating water molecule in the AMSH-LP^{E292A}•K63-Ub₂ complex. Nonetheless, the active site architecture in the K63-Ub₂-bound AMSH-LP^{E292A} is nearly identical to that in AMSH-LP alone. All atoms of the Zn^{2+} -coordinating residues in AMSH-LP can be superposed on those of the K63-Ub₂-bound AMSH-LP^{E292A} with a root mean squared deviation value of 0.235 Å. On the basis of this superposition, we have modelled Zn^{2+} , the Zn^{2+} -coordinating water molecule, and the side chain of Glu 292 in the AMSH-LP^{E292A}•K63-Ub₂ complex shown in Figs 3c and 4a.

The catalytic mechanisms of Zn^{2+} -containing metalloproteases have been extensively studied^{30,31}. The intermediate state of thermolysin has been observed in structures of thermolysin bound to various phosphorous-containing peptide analogues³². Although AMSH-LP and thermolysin are only related by the fact that they are Zn^{2+} -containing metalloproteases, the Zn^{2+} -coordinating active sites are remarkably similar to each other (Fig. 4a, b). When substrate-bound structures of thermolysin and AMSH-LP are superposed so that the positions of the catalytic Zn^{2+} atoms and their coordinating atoms are aligned, the scissile bonds of the substrates are located in very similar locations and orientations (Fig. 4a, b). This suggests that the catalytic mechanism of isopeptide bond cleavage by AMSH family members is probably the same as that of peptide bond cleavage by thermolysin (Fig. 4c). We used our DUB assay to test whether mutants of the water-activating Glu 292, the Zn^{2+} -coordinating Asp 360, and the intermediate-stabilizing Ser 357 affect catalysis in a manner consistent with our proposed mechanism. We found that the E292A and D360A mutants completely lose their DUB activities and that the S357A mutant reduced the k_{cat} value 34-fold (Supplementary Table 1), suggesting that AMSH-LP probably catalyses isopeptide bond cleavage by means of the same mechanism that thermolysin uses to cleave peptide bonds.

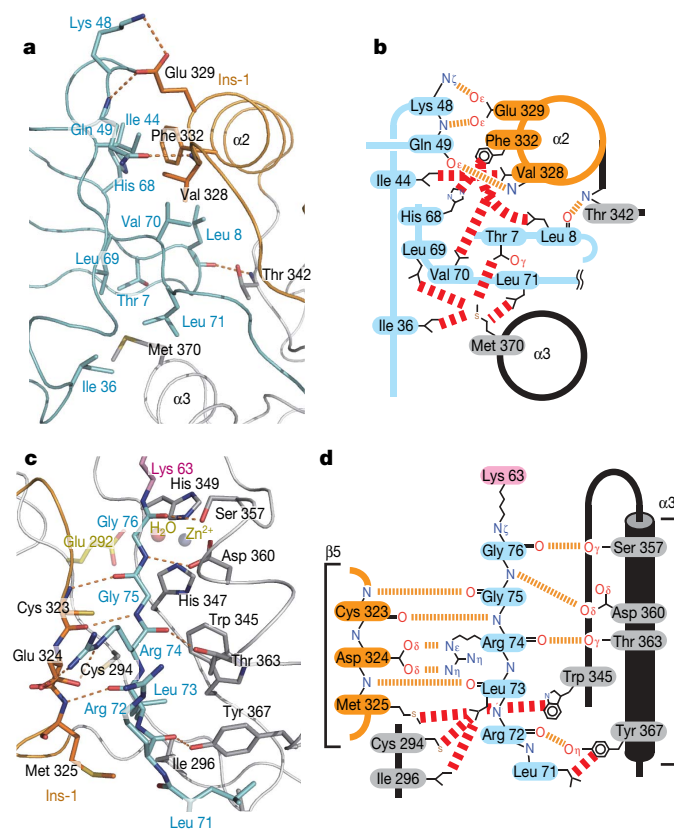


Figure 3 | Distal ubiquitin recognition. **a**, The hydrophobic interfaces between AMSH-LP and the distal ubiquitin. The colouring scheme is the same as in Fig. 1. **b**, Schematic representation of **a**. The colouring scheme is the same as in Fig. 2b, except that Ins-1 in AMSH-LP is highlighted by an orange background. **c**, The interface between AMSH-LP and the C-terminal tail of the distal ubiquitin. The colouring scheme is the same as in Fig. 1, except that the modelled side chain of Glu 292, Zn^{2+} and Zn^{2+} -coordinating water are translucent. **d**, Schematic representation of **c**. The colouring scheme is the same as in **b**.

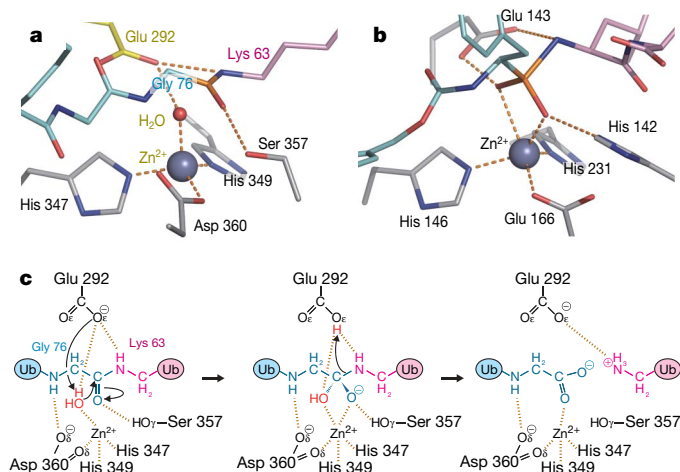


Figure 4 | Catalytic mechanism. **a**, The isopeptide bond of K63-Ub₂ in the active site of AMSH-LP. The colouring scheme is the same as in Fig. 3c, except that the carbon atom in the scissile bond is coloured orange. **b**, The reaction intermediate analogue CBZ-Phe-PO₂-Leu-Ala (ZFPLA) in the active site of thermolysin (protein data bank code 4TMN; CBZ is a benzyloxycarbonyl group). Thermolysin is coloured grey. CBZ-Phe, Leu-Ala and the phosphorous atom in ZEPLA are coloured cyan, pink and orange, respectively. **c**, Schematic representation of a proposed mechanism of DUB catalysis by AMSH-LP. The Zn²⁺-coordinating water and the distal and proximal ubiquitins are highlighted by red, cyan and pink, respectively.

Other JAMM isopeptidases

The human genome seems to contain 14 JAMM proteins, 7 of which have a complete set of the conserved residues for Zn²⁺ coordination⁷. Among them, six members (AMSH, AMSH-LP, BRCC36 (also known as BRCC3), Rpn11 (also known as PSMD14 and POH1), MYSM1 and CSN5) have been reported to have isopeptidase activity for ubiquitin or ubiquitin-like proteins^{7,33,34}. Although the amino acid sequences of these six members are highly divergent, portions of the JAMM core are conserved³⁵ (Supplementary Fig. 6). BRCC36, Rpn11, MYSM1 and CSN5 have insertions in a similar position of AMSH-family Ins-1, which might provide the β -sheet interaction with the C-terminal tail of the distal ubiquitin or ubiquitin-like proteins.

Rpn11 is a component of the 19S proteasome and cleaves the isopeptide bond of the polyubiquitin chain proximal to the substrate before substrate degradation by the proteasome^{36,37}. Alignment of the amino acid sequences of AMSH, AMSH-LP and Rpn11 suggests that Rpn11 uses an AMSH-like strategy to bind distal ubiquitin. The sequence alignment of Rpn11 and AMSH-LP shows that two insertions intervene in the JAMM core of Rpn11 as well as in that of AMSH-LP (Supplementary Fig. 7). The JAMM motif and the distal ubiquitin-binding residues are well conserved between AMSH-LP and Rpn11. Furthermore, residues Phe 332 and Val 328 are involved in distal ubiquitin binding by AMSH-LP and are strictly conserved in Rpn11, whereas Glu 329 is replaced by the functionally equivalent Asp (Supplementary Fig. 7). This sequence conservation suggests that AMSH family members and Rpn11 share a common strategy for distal ubiquitin recognition and DUB catalysis (Fig. 5). In marked contrast, there is no sequence similarity between Rpn11 and AMSH-LP around the proximal ubiquitin-binding residues. Furthermore, Ins-2 of Rpn11 lacks a Zn²⁺-coordinating motif, which we demonstrated is crucial for AMSH-LP DUB catalysis owing to stabilization of the proximal ubiquitin recognition loop in Ins-2 (Supplementary Fig. 7).

The observations discussed above suggest that the specificity of AMSH family members for Lys 63-linked polyubiquitin chains is derived primarily from the interactions with the proximal ubiquitin, which we revealed by the AMSH-LP^{E292A}•K63-Ub₂ complex structure in this study. In AMSH-LP, the linkage-specific recognition of the polyubiquitin chain is achieved by a single domain containing

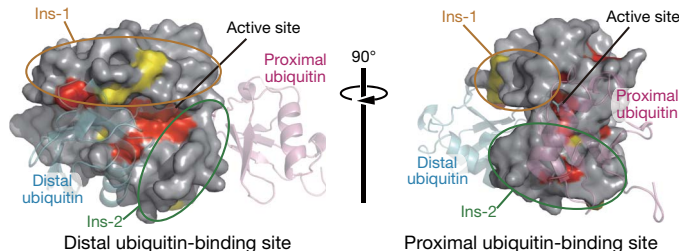


Figure 5 | Conservation among the DUB domains of AMSH, AMSH-LP and Rpn11. 100% and more than 60% identical residues are highlighted by red and yellow, respectively, on the molecular surface of AMSH-LP. K63-Ub₂ is shown as a translucent cartoon model with the proximal and distal ubiquitins coloured pink and cyan, respectively.

two characteristic insertions. However, many DUBs have a multi-domain structure, and some DUBs associate with other proteins, including E3 ubiquitin ligases⁷. It is probable that, in some cases, these additional domains or proteins can confer linkage specificity. A better understanding of the mechanisms used by proteins to achieve linkage-specific polyubiquitin cleavage awaits the further biochemical and structural characterization of other linkage-specific DUBs.

METHODS SUMMARY

Details of the preparation and crystallization of the AMSH-LP DUB domain alone and in complex with K63-Ub₂ are described in Methods. Diffraction data sets were collected at beamline NW12A in PF-AR (Tsukuba, Japan), and processed with HKL2000 (ref. 38) and the CCP4 program suite³⁹. To solve the structure of the AMSH-LP DUB domain from the SAD data set, SHELX97 (ref. 40), SHARP⁴¹, SOLOMON/DM^{42,43} and ARP/wARP⁴⁴ were used for heavy-atom site search, phase calculation, density improvement and automatic model-building, respectively. These programs were controlled by autoSHARP⁴⁵. MOLREP⁴⁶ was used to solve the structure of the AMSH-LP^{E292A}•K63-Ub₂ complex by molecular replacement. Atomic models were corrected by using COOT⁴⁷ with careful inspection. Refinement was carried out by using Refmac5 (ref. 48) with iterative correction and refinement of the atomic models. The final models have excellent stereochemistry (except that His 303 of AMSH-LP in the complex is in the disallowed region of the Ramachandran plot) and R_{free} values of 0.165 for the AMSH-LP DUB domain and 0.215 for the AMSH-LP^{E292A}•K63-Ub₂ complex at 1.2 Å and 1.6 Å resolution, respectively. Data collection, phasing and refinement statistics are shown in Supplementary Table 2. All molecular graphics were prepared with PyMOL (DeLano Scientific; <http://www.pymol.org>).

For the *in vitro* DUB assay, wild-type and mutant AMSH-LP DUB domains were overproduced in *Escherichia coli* as glutathione S-transferase (GST)-fusion proteins and affinity-purified by glutathione sepharose chromatography. Site-directed mutations were generated by PCR. For each experiment, 1 μ g of enzyme was incubated with 10 μ g of substrate at 20 °C in 50 μ l of 100 mM Tris-HCl buffer (pH 7.2) containing 25 mM KCl, 5 mM MgCl₂ and 1 mM DTT. Reactions were stopped by boiling after addition of 2 \times SDS-PAGE sample loading buffer and then analysed by SDS-PAGE. Details of the kinetic experiments are described in Methods.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 2 May; accepted 15 July 2008.

Published online 31 August 2008.

- Glickman, M. H. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428 (2002).
- Hicke, L. Protein regulation by monoubiquitin. *Nature Rev. Mol. Cell Biol.* **2**, 195–201 (2001).
- Miranda, M. & Sorkin, A. Regulation of receptors and transporters by ubiquitination: new insights into surprisingly similar mechanisms. *Mol. Interv.* **7**, 157–167 (2007).
- Huang, F., Kirkpatrick, D., Jiang, X., Gygi, S. & Sorkin, A. Differential regulation of EGF receptor internalization and degradation by multiubiquitination within the kinase domain. *Mol. Cell* **21**, 737–748 (2006).
- Saksena, S., Sun, J., Chu, T. & Emr, S. D. ESCRTing proteins in the endocytic pathway. *Trends Biochem. Sci.* **32**, 561–573 (2007).

6. Amerik, A. Y. & Hochstrasser, M. Mechanism and function of deubiquitinating enzymes. *Biochim. Biophys. Acta* **1695**, 189–207 (2004).
7. Nijman, S. M. *et al.* A genomic and functional inventory of deubiquitinating enzymes. *Cell* **123**, 773–786 (2005).
8. Komada, M. Controlling receptor downregulation by ubiquitination and deubiquitination. *Curr. Drug Discov. Technol.* **5**, 78–84 (2008).
9. Clague, M. J. & Urbe, S. Endocytosis: the DUB version. *Trends Cell Biol.* **16**, 551–559 (2006).
10. Mizuno, E. *et al.* Regulation of epidermal growth factor receptor down-regulation by UBPY-mediated deubiquitination at endosomes. *Mol. Biol. Cell* **16**, 5163–5174 (2005).
11. McCullough, J., Clague, M. J. & Urbe, S. AMSH is an endosome-associated ubiquitin isopeptidase. *J. Cell Biol.* **166**, 487–492 (2004).
12. Kim, M. S., Kim, J. A., Song, H. K. & Jeon, H. STAM–AMSH interaction facilitates the deubiquitination activity in the C-terminal AMSH. *Biochem. Biophys. Res. Commun.* **351**, 612–618 (2006).
13. Hu, M. *et al.* Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* **111**, 1041–1054 (2002).
14. Reyes-Turcu, F. E. *et al.* The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. *Cell* **124**, 1197–1208 (2006).
15. Misaghi, S. *et al.* Structure of the ubiquitin hydrolase UCH-L3 complexed with a suicide substrate. *J. Biol. Chem.* **280**, 1512–1520 (2005).
16. Messick, T. E. *et al.* Structural basis for ubiquitin recognition by the OTU1 ovarian tumor domain protein. *J. Biol. Chem.* **283**, 11038–11049 (2008).
17. Varadan, R., Assfalg, M., Raasi, S., Pickart, C. & Fushman, D. Structural determinants for selective recognition of a Lys48-linked polyubiquitin chain by a UBA domain. *Mol. Cell* **18**, 687–698 (2005).
18. Trempe, J. F. *et al.* Mechanism of Lys48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *EMBO J.* **24**, 3178–3189 (2005).
19. Eddins, M. J., Carlile, C. M., Gomez, K. M., Pickart, C. M. & Wolberger, C. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. *Nature Struct. Mol. Biol.* **13**, 915–920 (2006).
20. Pena, V., Liu, S., Bujnicki, J. M., Luhrmann, R. & Wahl, M. C. Structure of a multipartite protein–protein interaction domain in splicing factor Prp8 and its link to *Retinitis pigmentosa*. *Mol. Cell* **25**, 615–624 (2007).
21. Zhang, L. *et al.* Crystal structure of the C-terminal domain of splicing factor Prp8 carrying *Retinitis pigmentosa* mutants. *Protein Sci.* **16**, 1024–1031 (2007).
22. Sanches, M., Alves, B. S., Zanchin, N. I. & Guimaraes, B. G. The crystal structure of the human Mov34 MPN domain reveals a metal-free dimer. *J. Mol. Biol.* **370**, 846–855 (2007).
23. Ambroggio, X. I., Rees, D. C. & Deshaies, R. J. JAMM: a metalloprotease-like zinc site in the proteasome and signalosome. *PLoS Biol.* **2**, E2 (2004).
24. Tran, H. J., Allen, M. D., Lowe, J. & Bycroft, M. Structure of the Jab1/MPN domain and its implications for proteasome function. *Biochemistry* **42**, 11460–11465 (2003).
25. Pickart, C. M. & Raasi, S. Controlled synthesis of polyubiquitin chains. *Methods Enzymol.* **399**, 21–36 (2005).
26. Hofmann, R. M. & Pickart, C. M. *In vitro* assembly and recognition of Lys-63 polyubiquitin chains. *J. Biol. Chem.* **276**, 27936–27943 (2001).
27. Tenno, T. *et al.* Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. *Genes Cells* **9**, 865–875 (2004).
28. Varadan, R. *et al.* Solution conformation of Lys63-linked di-ubiquitin chain provides clues to functional diversity of polyubiquitin signaling. *J. Biol. Chem.* **279**, 7055–7063 (2004).
29. Hurley, J. H., Lee, S. & Prag, G. Ubiquitin-binding domains. *Biochem. J.* **399**, 361–372 (2006).
30. Gupta, S. P. Quantitative structure–activity relationship studies on zinc-containing metalloproteinase inhibitors. *Chem. Rev.* **107**, 3042–3087 (2007).
31. Lipscomb, W. N. & Strater, N. Recent advances in zinc enzymology. *Chem. Rev.* **96**, 2375–2434 (1996).
32. Holden, H. M., Tronrud, D. E., Monzingo, A. F., Weaver, L. H. & Matthews, B. W. Slow- and fast-binding inhibitors of thermolysin display different modes of binding: crystallographic analysis of extended phosphoramidate transition-state analogues. *Biochemistry* **26**, 8542–8553 (1987).
33. Zhu, P. *et al.* A histone H2A deubiquitinase complex coordinating histone acetylation and H1 dissociation in transcriptional regulation. *Mol. Cell* **27**, 609–621 (2007).
34. Sobhian, B. *et al.* RAP80 targets BRCA1 to specific ubiquitin structures at DNA damage sites. *Science* **316**, 1198–1202 (2007).
35. Maytal-Kivity, V., Reis, N., Hofmann, K. & Glickman, M. H. MPN⁺, a putative catalytic motif found in a subset of MPN domain proteins from eukaryotes and prokaryotes, is critical for Rpn11 function. *BMC Biochem.* **3**, 28 (2002).
36. Yao, T. & Cohen, R. E. A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* **419**, 403–407 (2002).
37. Verma, R. *et al.* Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* **298**, 611–615 (2002).
38. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
39. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
40. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. D* **64**, 112–122 (2008).
41. de la Fortelle, E. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
42. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52**, 30–42 (1996).
43. Cowtan, K. D. & Main, P. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr. D* **49**, 148–157 (1993).
44. Morris, R. J., Perrakis, A. & Lamzin, V. S. ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.* **374**, 229–244 (2003).
45. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230 (2007).
46. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Cryst.* **30**, 1022–1025 (1997).
47. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
48. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Toyoshima for support of this research. We thank M. Shirakawa and T. Tenno for advice about the K63–Ub₂ preparation. We thank S. Kaiser for critical reading and improvement of this manuscript. We thank the beam-line staffs at NW12A of PF-AR (Tsukuba, Japan) for technical help during data collection. This work was supported by grants from MEXT to S.F. and O.N. Y.S. is supported by JSPS research fellowships for young scientists.

Author Contributions Y.S. carried out the crystallization and structure determination of the AMSH-LP DUB domain and its complex with K63–Ub₂. A.Y. prepared K63–Ub₂ and polyubiquitin chains with support from K.O., M.Y., O.N., K.I. and M.K. Y.S. and A.Y. carried out the *in vitro* DUB assays. K.I. and M.K. provided resources for K63–Ub₂ preparation. A.Y., H.M. and S.F. assisted with the crystallization and structure determination. Y.S. and S.F. wrote the paper, with editing from A.Y., H.M., O.N., K.I. and M.K. S.F. and M.K. designed the research. All authors discussed the results and commented on the manuscript. S.F. supervised the work.

Author Information The coordinates and structure factors of the AMSH-LP DUB domain and its complex with K63–Ub₂ have been deposited in the Protein Data Bank with the accession codes 2ZNR and 2ZNV, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.F. (fukai@iam.u-tokyo.ac.jp).

METHODS

Preparation of the AMSH-LP DUB domain. The gene encoding the DUB domain of human AMSH-LP (residues 264–436) was PCR-amplified and cloned into the expression vector pCold-GST (Hayashi, K. and Kojima, C., submitted) with NdeI and XhoI sites to produce an N-terminal GST fusion protein. *E. coli* Rosetta (DE3) cells (Invitrogen) were transformed with the expression vector and cultured in luria broth media containing 100 mg l⁻¹ ampicillin at 37 °C. When the optical density at 600 nm of the culture reached ~0.5, the culture was incubated for 30 min at 15 °C. Thereafter, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.3 mM to induce protein expression for 24 h at 15 °C. Cells were collected by centrifugation at 8,000g for 15 min, and were disrupted by sonication in phosphate buffered saline (PBS) containing 1 mM dithiothreitol (DTT), 1 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5% Triton X-100. The lysates were centrifuged at 30,000g for 60 min, and the supernatants were loaded onto a glutathione sepharose FF column (GE Healthcare) pre-equilibrated with PBS containing 1 mM DTT and 0.5% Triton X-100. The column was washed with PBS containing 1 mM DTT and 0.5% Triton X-100 and then with PBS containing 1 mM DTT. The GST fusion proteins were eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 200 mM NaCl, 1 mM DTT and 15 mM reduced glutathione. The GST tags were cleaved by PreScission protease (GE Healthcare), and the samples were dialysed against 50 mM Tris-HCl buffer (pH 8.0) containing 40 mM NaCl and 1 mM DTT. The samples were passed over a ResourceQ anion exchange column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl buffer (pH 8.0) containing 40 mM NaCl and 1 mM DTT. The flow-through fractions were loaded onto a Superdex 200 16/60 (prep grade) column (GE Healthcare) pre-equilibrated with 10 mM Tris-HCl buffer (pH 7.2) containing 50 mM NaCl and 5 mM β-mercaptoethanol.

Preparation of mouse E1. The N-terminal His₆-tagged mouse E1 (Ube1, also known as Uba2) was overproduced in Sf9 cells infected by baculovirus containing the His₆-tagged mouse E1 gene⁴⁹. Cells were collected and disrupted by sonication in 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl, 5 mM β-mercaptoethanol and 0.1 mM PMSF. The cleared lysate was loaded onto Ni-NTA (Qiagen) column. The column was washed with 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl, 5 mM β-mercaptoethanol and 20 mM imidazole. E1 was eluted with a step gradient of 50–800 mM imidazole. The eluted E1 was dialysed against 50 mM Tris-HCl buffer (pH 7.5) containing 150 mM NaCl and 5 mM β-mercaptoethanol. Finally, the His₆-tagged E1 was concentrated to 4 μM and stored at -80 °C until use.

Preparation of yeast E2, Ubc13 and Mms2. The genes encoding yeast Ubc13 and Mms2 were cloned into pGEX-4T1 (GE Healthcare) and pET28a (Novagen) expression vectors, respectively. *E. coli* strain Rosetta (DE3) cells (Invitrogen) were transformed with each expression vector to produce the N-terminal GST-fused Ubc13 or His₆-tagged Mms2. The transformed *E. coli* cells were cultured in luria broth media containing 50 mg l⁻¹ ampicillin for Ubc13 or kanamycin for Mms2 at 37 °C. When the optical density at 600 nm of the culture reached ~0.5, IPTG was added to a final concentration of 0.3 mM to induce protein expression. After another 17-h culture at 20 °C, cells were collected by centrifugation at 8,000g for 15 min, and were disrupted by sonication in PBS containing 1 mM DTT and 1 mM PMSF. The His₆-tagged Mms2 was purified in the same manner as the His₆-tagged E1. The purified Mms2 was concentrated to 140 μM and stored at -80 °C until use. The cleared lysate containing the GST-fused Ubc13 was loaded onto a glutathione sepharose FF column (GE Healthcare). The column was washed with PBS containing 1 mM DTT. The GST-fused Ubc13 was eluted with 50 mM Tris-HCl (pH 8.0) buffer containing 150 mM NaCl, 15 mM reduced glutathione and 1 mM DTT, and cleaved by thrombin at 4 °C overnight (16h) to remove the GST tag. After dialysis against 50 mM Tris-HCl (pH 8.0) buffer containing 50 mM NaCl and 1 mM DTT, Ubc13 was subjected to MonoQ anion exchange column (GE Healthcare) pre-equilibrated with 50 mM Tris-HCl buffer (pH 8.0) containing 1 mM DTT. Ubc13 was eluted with 50 mM Tris-HCl buffer (pH 8.0) containing 1 mM DTT using a linear gradient of 0–1.0 M NaCl. The eluted Ubc13 was concentrated to 130 μM and stored at -80 °C until use.

Preparation of mouse ubiquitin mutants, K63R and D77. The genes encoding mouse ubiquitin mutants K63R and D77 were cloned into pET26b (Novagen) expression vector. *E. coli* strain Rosetta (DE3) cells (Invitrogen) were transformed with the expression vector to produce the ubiquitin mutants. The transformed *E. coli* cells were cultured in luria broth media containing 50 mg l⁻¹ kanamycin at 37 °C. When the optical density at 600 nm of the culture reached ~0.5, IPTG was added to a final concentration of 0.3 mM to induce protein expression. After expression for 17 h at 20 °C, cells were collected by centrifugation at 8,000g for 15 min, and were disrupted by sonication in 50 mM ammonium acetate (pH 4.5) buffer containing 2 mM DTT, 1 mM EDTA and 1 mM PMSF. The cleared lysate was incubated at 70 °C for 5 min to denature proteins from *E. coli*, which were precipitated by centrifugation at 30,000g for 20 min. The

supernatant was loaded onto SP Sepharose Fast Flow (GE Healthcare) pre-equilibrated with 50 mM ammonium acetate (pH 4.5) buffer containing 2 mM DTT and 0.1 mM EDTA. The ubiquitin mutants were eluted with 50 mM ammonium acetate (pH 4.5) buffer containing 2 mM DTT using a linear gradient of 0–500 mM NaCl. The eluted ubiquitin mutants were fractionated by HiLoad Superdex75 (GE Healthcare) size-exclusion column with 50 mM Tris-HCl (pH 7.6) buffer containing 2 mM DTT and 0.1 mM EDTA. Each mutant was concentrated to 30 g l⁻¹ and stored at -80 °C until use.

K63-Ub₂ synthesis and purification of the complex. Purified mouse E1, yeast E2 (Ubc13 and Mms2) and two mouse ubiquitin mutants (K63R and D77) were mixed in 50 mM Tris-HCl buffer (pH 7.6) containing 2 mM ATP, 1 mM DTT, 5 mM MgCl₂, 10 mM creatine phosphate (Sigma Aldrich), 0.6 units of creatine phosphokinase (Sigma Aldrich) and 0.6 units of inorganic pyrophosphatase (Sigma Aldrich) at 37 °C overnight. Final concentrations of E1, E2 and ubiquitin mutants were 0.1 μM, 8 μM and 5 g l⁻¹, respectively. The reaction mixture was put on ice for 10 min to stop the reaction. 0.2 volumes of 2 M ammonium acetate were added to decrease the pH to less than 4.0. The synthesized K63-Ub₂ was loaded onto a MonoS cation exchange column (GE Healthcare) pre-equilibrated with 50 mM ammonium acetate (pH 4.5) buffer containing 5 mM DTT and 1 mM EDTA. The K63-Ub₂ was eluted with 50 mM ammonium acetate (pH 4.5) buffer containing 5 mM DTT and 1 mM EDTA using a linear gradient of 0–0.6 M NaCl. The eluted K63-Ub₂ was loaded onto HiLoad Superdex75 (GE Healthcare) size-exclusion column with 20 mM phosphate-K (pH 7.6) buffer containing 150 mM KCl and 1 mM EDTA. The purified K63-Ub₂ was concentrated to 30 g l⁻¹ and stored at -80 °C until use.

To prepare the AMSH-LP^{E292A}•K63-Ub₂ complex, K63-Ub₂ was incubated with a 1.5-fold molar excess of AMSH-LP^{E292A} (residues 264–436) at 4 °C overnight. The AMSH-LP^{E292A}•K63-Ub₂ complex was loaded onto a HiLoad Superdex 200 size-exclusion column (GE Healthcare) equilibrated with 10 mM Tris-HCl buffer (pH 7.2), 50 mM NaCl and 5 mM β-mercaptoethanol to remove unbound AMSH-LP^{E292A}.

Crystallization. Purified AMSH-LP (residues 264–436) and the AMSH-LP^{E292A}•K63-Ub₂ complex were concentrated to 13.5 and 11.5 mg ml⁻¹, respectively, using an Amicon Ultra-15 10,000 MWCO filter (Millipore). Initial crystallization screening was performed using the sitting drop vapour diffusion method at 20 °C, with a Mosquito liquid-handling robot (TTP Lab Tech). We tested about 500 conditions, using crystallization reagent kits supplied by Hampton Research, and initial hits were further optimized. The best crystals of AMSH-LP were obtained at 20 °C with the sitting drop vapour diffusion method by mixing 1 μl of protein solution with an equal volume of precipitant solution containing 45 mM sodium acetate buffer (pH 4.6), 22% PEG4000, 90 mM ammonium acetate and 10 mM praseodymium (III) acetate and were allowed to equilibrate against 500 μl of reservoir solution containing 50 mM sodium acetate buffer (pH 4.6), 18% PEG4000 and 100 mM ammonium acetate. The best crystals of the AMSH-LP^{E292A}•K63-Ub₂ complex were obtained at 20 °C with the sitting drop vapour diffusion method by mixing 1 μl of protein solution with an equal amount of precipitant solution containing 180 mM tri-ammonium citrate (pH 7.0), 24% PEG3350 and 3% 1,6-hexanediol and allowed to equilibrate against 500 μl of reservoir solution containing 200 mM tri-ammonium citrate (pH 7.0) and 20% PEG3350. The crystal of the AMSH-LP belonged to space group P6₅, with unit cell dimensions $a = b = 81.9 \text{ \AA}$, $c = 64.7 \text{ \AA}$. The crystal of the AMSH-LP^{E292A}•K63-Ub₂ complex belonged to space group P2₁, with unit cell dimensions $a = 38.1 \text{ \AA}$, $b = 97.4 \text{ \AA}$, $c = 87.9 \text{ \AA}$, $\beta = 97.5^\circ$.

Kinetic analysis. Kinetic analyses were performed at 20 °C in a reaction buffer containing 100 mM Tris-HCl buffer (pH 7.2), 25 mM KCl, 5 mM MgCl₂ and 1 mM DTT. The final concentrations of the AMSH-LP DUB domain in all reaction mixtures were 25 nM for the wild type and 50 (E329A, F332A, M370A), 250 (S358A), 500 (T353A, S357A, F407A) or 1,000 nM (F355A, C402S) for the mutants. To calculate the kinetic parameters, initial rates were measured at four different concentrations of K63-Ub₂ ranging from 25 to 1,200 μM (Supplementary Fig. 8). The reaction was terminated by boiling after addition of 2× SDS-PAGE sample loading buffer then subjected to tricine SDS-PAGE analysis. Various amounts (4.3 μg to 12.8 μg) of ubiquitin monomers were loaded on the same gels as standards for quantification, and stained with coomassie brilliant blue. Gels were imaged using a Nikon COOLPIX 995 digital camera (Nikon Corporation) in 8-bit TIFF format of 2,048 × 1,536 pixels (typical sample images are presented in Supplementary Fig. 8). The pixel density of the gel bands of ubiquitin monomers was analysed with ImageJ software (<http://rsb.info.nih.gov/ij/>). The assay was carried out three times for each mutant. Kinetic parameters were calculated by fitting to a Hanes-Woolf plot.

49. Iwai, K. *et al.* Identification of the von Hippel–Lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex. *Proc. Natl Acad. Sci. USA* **96**, 12436–12441 (1999).

Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes

Dong-Sun Lee^{1*}, Pierre Nioche^{2*}, Mats Hamberg³ & C. S. Raman¹

The oxylipin pathway generates not only prostaglandin-like jasmonates but also green leaf volatiles (GLVs), which confer characteristic aromas to fruits and vegetables. Although allene oxide synthase (AOS) and hydroperoxide lyase are atypical cytochrome P450 family members involved in the synthesis of jasmonates and GLVs, respectively, it is unknown how these enzymes rearrange their hydroperoxide substrates into different products. Here we present the crystal structures of *Arabidopsis thaliana* AOS, free and in complex with substrate or intermediate analogues. The structures reveal an unusual active site poised to control the reactivity of an epoxyallylic radical and its cation by means of interactions with an aromatic π -system. Replacing the amino acid involved in these steps by a non-polar residue markedly reduces AOS activity and, unexpectedly, is both necessary and sufficient for converting AOS into a GLV biosynthetic enzyme. Furthermore, by combining our structural data with bioinformatic and biochemical analyses, we have discovered previously unknown hydroperoxide lyase in plant growth-promoting rhizobacteria, AOS in coral, and epoxyalcohol synthase in amphioxus. These results indicate that oxylipin biosynthetic genes were present in the last common ancestor of plants and animals, but were subsequently lost in all metazoan lineages except Placozoa, Cnidaria and Cephalochordata.

Oxylipins are bioactive lipids derived from oxygenation of polyunsaturated fatty acids. Both jasmonates^{1–4} and GLVs^{5,6} are representative plant oxylipins. Whereas jasmonates are essential for plant development and host immunity^{7–9}, GLVs are released to counter biotic and abiotic stresses, such as wounding, herbivore attacks and ozone exposure¹⁰. Curiously, jasmonates and animal prostaglandins^{11,12} share notable structural and functional properties (Supplementary Figs 1 and 2).

AOS^{13–18} catalyses the first committed step¹⁹ in the synthesis of jasmonates, and hydroperoxide lyase (HPL)^{5,20,21} generates GLVs (Fig. 1). Some plants use divinyl ether synthase (DES) to produce divinyl ethers^{22,23}, which are thought to have a role in host defense. AOS, HPL, and DES belong to the cytochrome P450 (P450 or CYP) superfamily^{16,24}, which currently has more than 8,700 members distributed across all three domains of life. All P450 enzymes are thiolate-coordinated haem proteins and most of them function as monooxygenases²⁵. Although there is <20% sequence identity within the superfamily, P450s share a common protein fold, as shown by nearly 30 unique crystal structures of bacterial and human enzymes^{26,27}. However, there is no structural information on plant P450s.

AOS, HPL and DES are members of the CYP74 family. These unusual P450s do not function as monooxygenases, but instead rearrange fatty acid hydroperoxides into structurally different products (Fig. 1). Despite four decades of research, several observations concerning CYP74 biochemistry and catalysis have remained enigmatic. Here we present the crystal structure of *A. thaliana* AOS (At-AOS, also known as CYP74A) and illustrate how an important new activity evolved. We also show that a single substitution changes the product specificity of AOS to that of an HPL. Finally, we report that CYP74 enzymes are not confined to plants, but also occur in bacteria and animals.

Overall structure of AOS

At-AOS adopts the characteristic fold of P450s (Fig. 2a, Supplementary Tables 1 and 2, and Supplementary Fig. 3). The membrane-binding

region can be identified by bound detergent molecules and by its rather extensive non-polar surface ($\sim 2,400 \text{ \AA}^2$; Fig. 2a and Supplementary

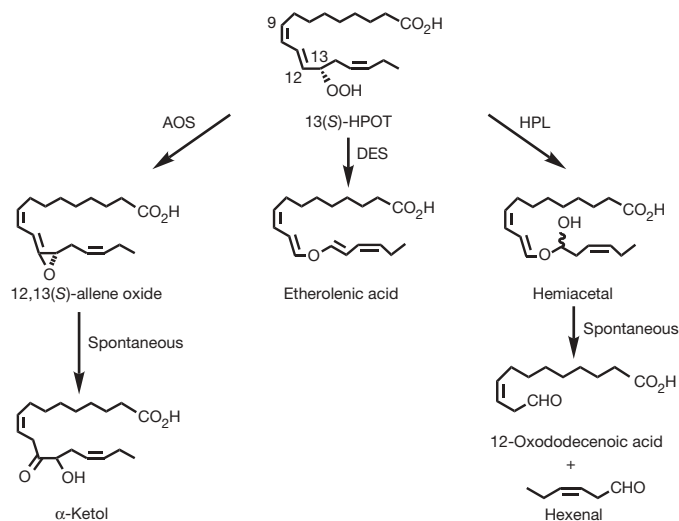


Figure 1 | Reactions catalysed by the CYP74 enzyme family. In higher plants, C₁₈ fatty acids (linoleic and linolenic acids) are oxygenated at either position 9 or 13 by lipoxygenases to yield hydroperoxides. Subsequently, these are converted by allene oxide synthase (AOS, also known as CYP74A), hydroperoxide lyase (HPL, also known as CYP74B) and divinyl ether synthase (DES, also known as CYP74D) to allene oxide (an essential intermediate in jasmonate biosynthesis), green leaf volatiles (aldehydes) and divinyl ethers, respectively. For clarity, only 13-hydroperoxide-derived metabolites are shown. 12,13(S)-allene oxide, 12,13S-epoxy-9Z,11,15Z-octadecatrienoic acid; 13(S)-HPOT, 13S-hydroperoxyoctadecatrienoic acid; α -ketol is the hydrolytic product of the highly unstable allene oxide; hemiacetal, hydroxyhexenyloxydodecadienoic acid.

¹Department of Biochemistry and Molecular Biology, University of Texas Medical School, Houston, Texas 77030, USA. ²Université Paris Descartes, INSERM UMR-S 747, 75270 Paris cedex 06, France. ³Division of Physiological Chemistry II, Department of Medical Biochemistry & Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden.

*These authors contributed equally to this work.

Fig. 4a, b). It also harbours the entry point for the 22-Å deep substrate access channel (Supplementary Fig. 4c). AOS lacks the characteristic amino-terminal transmembrane domain of microsomal P450s²⁸ and, yet, associates very tightly with chloroplastic envelope membranes²⁹ and plastoglobules³⁰. A comparison of the crystal structures of At-AOS and mammalian CYP2C5 (ref. 28) reveals that both P450s use the same macromolecular surface to interact with the membrane (Supplementary Fig. 5).

The haem cofactor of At-AOS is inserted between the I-helix and the L-helix and, like classical P450s^{26,27}, its propionate groups interact with strictly conserved basic residues through hydrogen bonds (Fig. 2b). Haem is in the low-spin state, with Cys471 and a water molecule serving as proximal and distal coordination ligands of Fe(III), respectively (Fig. 2b and Supplementary Fig. 6a). However, there are three major changes that distinguish At-AOS from all other known P450 structures. First, nine residues are inserted into the middle of the signature haem-binding loop harbouring the Cys471 ligand (Fig. 2c, Supplementary Fig. 6b and Supplementary Discussion). This is extraordinary because even a single residue insertion within this region is not observed in >8,000 P450s. This modification, which is preserved in all CYP74 family members, not only alters the proximal site topology but also causes extensive remodelling of the macromolecular surface essential for redox partner interactions in typical P450s (Supplementary Fig. 6c). In addition, the Fe–S bond length is reproducibly longer (2.4 Å versus 2.2 Å) in the substrate-free state. Second, the kink centred on the second glycine of the I-helix (A/G)GxxT motif²⁶ in classical P450s is shifted by three residues towards the N terminus and is now localized on

Asn 321, positioning the amide group directly above the haem iron (4.4 Å; Fig. 2b and Supplementary Fig. 6d). Third, Ile 328 substitutes for the I-helix Thr that is critical for mediating proton delivery to haem-bound dioxygen in conventional P450s^{26,27}. Together, these structural modifications preclude monooxygenase activity in AOS. Notably, hydroperoxide-using CYP74 may have been an evolutionary intermediate during the transition from H₂O₂ (ref. 31) to O₂ as oxygen donor in P450 catalysis (Supplementary Figs 7–9 and Supplementary Discussion).

Identifying active site interactions

To gain insight into the structural basis of substrate recognition, we crystallized AOS together with 13(S)-HOT, a close mimic incorporating OH instead of the OOH group (Supplementary Fig. 10a). The structure shows well-defined electron density for both the fatty acid and the hydroxy portions of the molecule (Fig. 2d and Supplementary Video 1). 13(S)-HOT nicely conforms to the shape and physicochemical properties of the AOS active site. Its hydroxyl oxygen takes the place of a water molecule (wat1, Fig. 2b) in the substrate-free structure and interacts with the amide side chain of Asn 321 by a hydrogen bond. The carboxyl group engages in hydrogen bonding with Thr389, and the aliphatic segments maintain favourable hydrophobic contacts with neighbouring non-polar side chains. This binding mode explains why both diatomic (CO) and other axial haem ligands that attempt to coordinate the iron in a linear fashion experience difficulties due to the positioning of the carboxamide directly above the haem plane. Conversely, the peroxy group of the substrate approaches from the side and successfully ligates the haem while having a productive interaction with the catalytic Asn 321. Mutating Asn 321 abolishes >95% of the enzyme activity (Supplementary Fig. 10b).

Next we co-crystallized AOS with 12R,13S-vernolic acid, an analogue of the putative epoxylallylic reaction intermediate in AOS catalysis¹⁷ (Supplementary Fig. 10a). Our structure illustrates that the binding mode of 12R,13S-vernolic acid is very similar to that of 13(S)-HOT, with C11 residing in proximity to the aromatic face of Phe 137 (Fig. 3a and Supplementary Fig. 10c). Furthermore, our sequence analysis showed that whereas Phe 137 is strictly conserved in AOS, a Leu takes its place in HPL and DES (Supplementary Fig. 10d). Because enzymological studies¹⁷ have proposed that C11 is the site of both radical and carbocation formation, we hypothesized that removal of the electron-rich aromatic ring of Phe 137 would affect catalysis. Consistent with our prediction, At-AOS(F137L) is severely compromised in its ability to generate allene oxide (Fig. 3b). These observations identify an essential role for Phe 137. It is poised to stabilize the carbon-centred radical and to interact with the incipient carbocation by means of cation– π interactions (Supplementary Discussion).

Single point mutation confers HPL activity to AOS

We reasoned that, if the inability to stabilize a carbon-centred radical at C11 is the sole reason for loss of AOS activity, then the F137L mutant is likely to have altered product specificity. To test this, we probed for HPL and DES metabolites and found that At-AOS(F137L) exhibits robust HPL activity (Fig. 3b). However, DES products were not detected. To evaluate the molecular consequences of this mutation, we solved the structure of At-AOS(F137L) in complex with 12R,13S-vernolic acid (Fig. 3c and Supplementary Video 2) or with 13(S)-HOD (Supplementary Fig. 11a–d). The binding mode of both compounds is quite similar to that in the wild-type enzyme (Supplementary Fig. 11c, d). Notably, Phe has a bulkier side chain compared to Leu, but it is >5 Å away from the haem iron (Supplementary Fig. 11e), and therefore it is incapable of altering the reaction mechanism by means of steric effects.

Because our homology model of tomato HPL (Fig. 3c, grey overlay) and bioinformatic analysis showed that Ser 155 is substituted by an Ala in HPL (Supplementary Fig. 10d), we constructed the double mutant At-AOS(F137L,S155A) and found it to exhibit even stronger

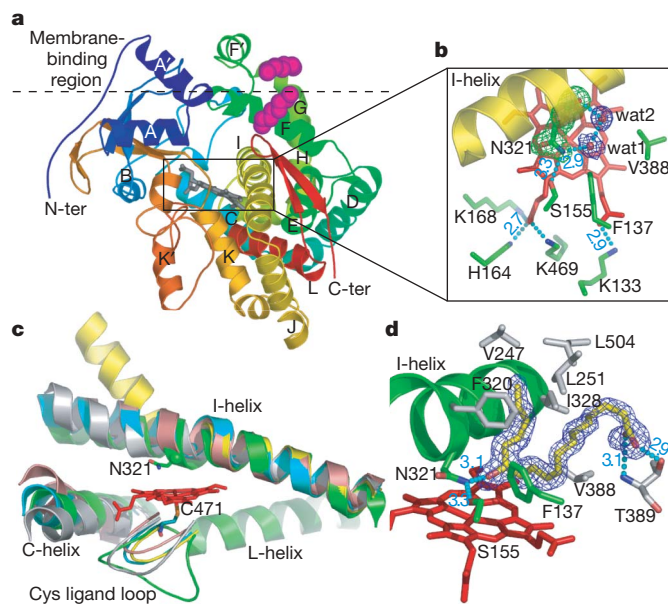


Figure 2 | Tertiary topology and substrate binding site of AOS. **a**, Ribbon diagram of At-AOS, rainbow-coloured from the N terminus (N-ter, blue) to the C terminus (C-ter, red). van der Waals representation (magenta) identifies hydrophobic tails of two detergent molecules interacting with F, F' and G helices, which are part of the membrane-associated nonpolar surface. Regions above and below the dashed line are facing the membrane and cytosol, respectively. **b**, Close-up view of the haem (red) environment. $F_o - F_c$ simulated annealing omit electron density (contoured at 3σ) is shown for Asn 321 (green) and two water molecules (wat1 and wat2, blue). **c**, Superposition of structural elements in AOS (green), CYP102A1 (pink, PDB ID: 1BVY), CYP3A4 (blue, PDB ID: 1TQN), CYP51 (yellow, PDB ID: 1EA1) and CYP152B1 (grey, PDB ID: 1IZO). **d**, Active site configuration. $F_o - F_c$ simulated annealing omit electron density (countoured at 3σ) of 13(S)-HOT (blue). Cyan dotted lines represent hydrogen bonds (donor–acceptor distances are in Å).

HPL activity. It converted the hydroperoxides 13(*S*)-HPOD, 13(*S*)-HPOT, 9(*S*)-HPOD and 11(*S*)-HPHT into the corresponding HPL products to a varying extent (Supplementary Fig. 12). Concerning the two first-mentioned hydroperoxides, 13(*S*)-HPOT gave a lower ratio of HPL/AOS products (approximately 25:75) than did 13(*S*)-HPOD (approximately 85:15). Structural analysis of the enzyme-bound 13(*S*)-HOT reveals that the C15 double bond is 3.3 Å away from C11, and it seems possible that the diminished effect of the F137L substitution on the formation of HPL products in the 13(*S*)-HPOT case can be attributed to a Phe-137-independent stabilization of the epoxyallylic intermediate by means of anchimeric assistance by π electrons from the C15=C16 double bond (Supplementary Fig. 13).

We also demonstrate that AOS to HPL conversion works across evolutionary boundaries. Although the eudicot (*Arabidopsis*) and monocot (*Oryza sativa*) branches of angiosperms diverged 140–180 million years ago³², introduction of F92L substitution into one of the two AOS genes found in rice (Supplementary Fig. 10d) readily changes product specificity (Supplementary Figs 14 and 15, and Supplementary Table 3).

Overall, we have established that AOS sandwiches C11 in between two π systems to ensure that 13(*S*)-HPOT, the only substrate that will yield jasmonates, is efficiently converted to allene oxide. Conversely, the HPL active site facilitates radical rearrangement by excluding a strategically positioned aromatic residue in the vicinity of C11 (Supplementary Discussion).

Structural mechanism for CYP74 catalysis

In both AOS and HPL, the terminal hydroperoxide oxygen of the substrate coordinates the ferric haem iron (2) (Fig. 4) after displacing

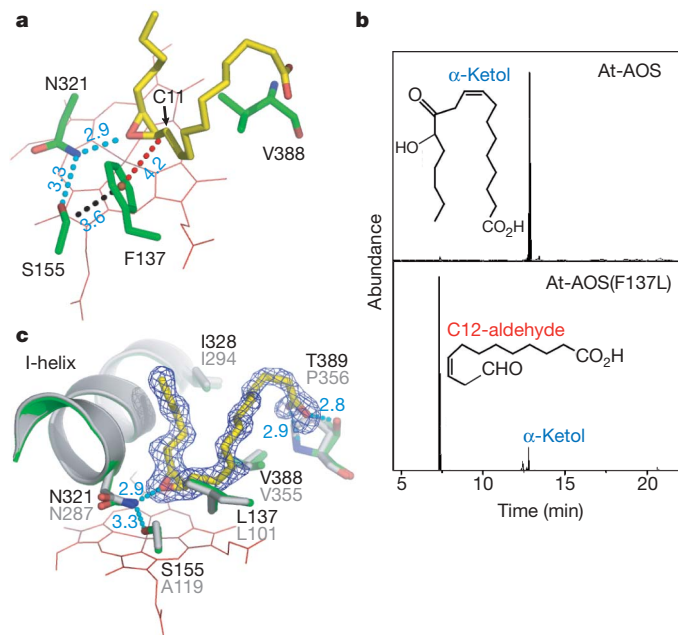


Figure 3 | Structural basis for evolving HPL activity from the AOS scaffold. **a**, At-AOS complexed with the reaction intermediate analogue (–)-vernolic acid (yellow). **b**, At-AOS(F137L) exhibits HPL activity. Recombinant wild-type and AOS(F137L) (0.2–4 μ M) were stirred with 13(*S*)-HPOD (200 μ M) at 23 °C for 15 min. After derivatization, product profiles were determined using gas chromatography–mass spectrometry (GC–MS) analysis with authentic compounds serving as references. The top panel shows At-AOS catalysed conversion to α -ketol (13-hydroxy-12-oxo-9*Z*-octadecenoic acid, *m/z* 412), the hydrolytic product of allene oxide. AOS(F137L) yielded 12-oxo-9*Z*-dodecenoic acid (C12-aldehyde, *m/z* 224), the product of HPL catalysis, as the major component (bottom panel). **c**, At-AOS(F137L) bound to (–)-vernolic acid (yellow) reveals the catalytic site structure of HPL. $F_o - F_c$ simulated annealing omit electron density, contoured at 3 σ , is shown in blue for the ligand. The homology model of tomato HPL (grey) is superimposed.

two bound water molecules (wat1 and wat2; see (1) and Fig. 2b) from the active site. With electrostatic assistance from a carboxamide (Supplementary Discussion), O–O homolysis³³ ensues, giving rise to an alkoxy radical (RO \cdot) and a protonated ferryl species (S–Fe(IV)–OH) (3) (Fig. 4). RO \cdot , while maintaining a hydrogen bond with the amide, adds to a proximate double bond (C11=C12) to generate an epoxide and a carbon-centred radical (C \cdot) at C11 (4). Hydrogen bonding to RO \cdot will favour cyclization over allylic hydrogen abstraction (Supplementary Discussion). The two mechanisms diverge at this stage. In the case of AOS, oxidation of

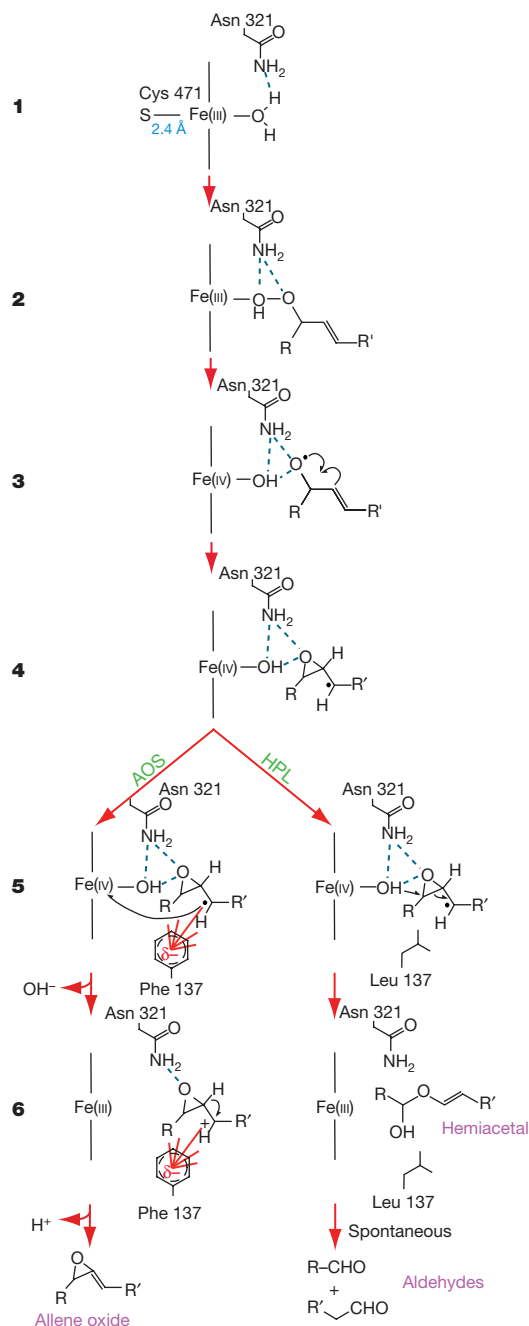


Figure 4 | Proposed reaction paths for AOS and HPL on the basis of the current structural and enzymological studies^{14–17,21}. The intermediate epoxyallylic radical formed in step 4 can either undergo one electron oxidation followed by proton loss (AOS) or oxygen rebound (HPL). The structure of the peroxide substrate is abbreviated to highlight the region undergoing chemical transformation. For clarity, the Fe–S bond between the haem iron and Cys 471 is only shown in step 1. It remains intact throughout the catalytic cycle. Hydrogen bonds are illustrated with blue dashed lines.

C[•] by S–Fe(IV)–OH (**5**, left) is made possible by Phe 137, which also stabilizes the incipient carbocation (C⁺) (at C11; **6**, left) by means of cation– π interactions. On β -proton elimination, a C=C bond is introduced adjacent to the epoxide, and allene oxide is generated (Supplementary Discussion). In the case of HPL there are two fundamental differences. First, electron transfer does not occur owing to difficulty in stabilizing the radical at this position. Consequently, the epoxycarbonyl radical undergoes C–C scission³⁴, yielding a radical at C13 that is able to oxygen rebound (iron-catalysed oxygen transfer involving radical recombination) (**5**, right) given the proximity to the ferryl centre (3.3 Å; Supplementary Fig. 11e). Second, the lack of a Phe 137 equivalent in HPL (**5**, right) prevents a carbocation intermediate from being adequately enriched or stabilized at C11. These conditions promote the formation of an unstable hemiacetal²¹, which spontaneously dissociates into short-chain aldehydes (**6**, right). In the case of DES, subtle alterations (Supplementary Fig. 16) may be sufficient for disabling oxygen rebound in favour of stereospecific hydrogen abstraction²⁴ by the protonated ferryl centre.

Although the CYP74 enzymes have evolved by precluding mono-oxygenation chemistry, they have used remarkable strategies to utilize the peroxide shunt pathway³³ and to exquisitely control the reactivity of catalytic intermediates for achieving product specificity. Nevertheless, our structural mechanism has both similarities and differences with the consensus hydroxylation mechanism of P450 monooxygenases³⁵. The protonated ferryl centre, S–Fe(IV)–OH (**3**), is a common intermediate in both systems. In the case of CYP74, a nine-residue insertion in the proximal Cys loop contributes significantly to diminishing the donor strength of the thiolate (Fig. 2c), favouring the generation of S–Fe(IV)–OH complex that can readily participate in electron transfer (AOS) or oxygen rebound (HPL). Conversely, for P450 monooxygenases³⁶, robust electron donation (push) from the proximal thiolate to the haem iron allows the ferryl-porphyrin cation radical (S–Fe(IV)=O + \bullet) to abstract a hydrogen atom from the substrate. The highly reactive S–Fe(IV)–OH (pK_a > 8.2) generated goes on to hydroxylate the substrate radical (Supplementary Discussion). The strong push effect of the thiolate also prevents S–Fe(IV)–OH from serving as a good electron acceptor³⁷.

Finally, we sought to compare the active sites of CYP74 and coral catalase-like AOS (cAOS)^{38,39} because of their ability to catalyse equivalent reactions (Supplementary Figs 1 and 17). We found that both systems control the reactivity of their haem cofactor in different ways to achieve the same end point (Supplementary Discussion).

Evolutionary history of oxylipin biosynthetic enzymes

On the one hand, cAOS has been found only in corals^{38,40} and in a cyanobacterium⁴¹ as a haem protein that is carboxy-terminally fused to a lipoxygenase (cAOS–LOX). On the other, the CYP74 family is

thought to be a plant invention⁴² because it has not been identified outside of angiosperms and moss²⁴. These observations raise three important questions. Is CYP74 unique to *Plantae*? Are these two allene oxide-generating protein families mutually exclusive in the same organism? And, could they have evolved independently and simultaneously?

Because several groups^{18,40} have failed to detect CYP74 in animals by means of homology-based polymerase chain reaction cloning, we took a structure-guided approach. By combining insights from At-AOS structure with multiple sequence alignments, we generated three unique motifs (Supplementary Table 4) for probing genome databases. Our search yielded more than 25 potential hits in marine invertebrates and 2 in rhizobacteria (Supplementary Fig. 18). Most of them show less than 30% sequence identity with the plant CYP74 family and fall short of the 40% identity required for being classified as *bona fide* CYP74 under the current P450 nomenclature system. However, they all retain the characteristic nine-residue insert in the proximal Cys loop, the amino acid makeup of which is notably similar to that found in plant CYP74 (Supplementary Figs 19 and 20). Next, we cloned and overexpressed CYP74-like proteins from *Methylobacterium nodulans* (Mn), *Acropora palmata* (Ap; Elkhorn coral) and *Branchiostoma floridae* (Bf; amphioxus) (Supplementary Figs 21–23). Mn-CYP74 and Ap-CYP74 show HPL and AOS activities, respectively (Fig. 5). Bf-CYP74 exhibits previously unknown epoxycarboxaldehyde synthase (EAS) activity by transforming the naturally occurring 13S-HPOD into a *cis*-epoxide product (Fig. 5). This is distinct from flaxseed AOS-catalysed¹⁷ conversion of an unnatural fatty acid hydroperoxide into an epoxycarboxaldehyde, where a *trans*-epoxide is generated. Although our homology models (Supplementary Fig. 24) show that the catalytic Phe is found in all three cases, we can surmise from the product profiles that it is suitably oriented for stabilizing the reaction intermediates only in Ap-AOS and Bf-EAS. The aromatic π -cloud is not in a position to influence the reaction path in Mn-HPL. This may also be the case with HPL from moss (Supplementary Fig. 18), which diverged from angiosperms over 400 million years ago⁴³. Thus, it is possible that an ancestral CYP74 may have functioned as a HPL before gene duplication and neofunctionalization.

Next, we examined the genetic context of CYP74 and uncovered five remarkable features. First, there are multiple CYP74 genes in animals. Amphioxus has 20, whereas *Nematostella* has just 2 (Fig. 6). This is also true in plants; grapevine and *Arabidopsis* encode 6 and 2 genes, respectively. Second, animal CYP74 genes contain numerous introns (Supplementary Table 5). In contrast, most of the plant CYP74 genes are intronless. Third, sea anemone and *Trichoplax* CYP74 genes propagate as tandem arrays (Supplementary Fig. 25). By querying the plant genomes, we discovered that grapevine, poplar and rice CYP74 genes also do the same. Fourth, a LH2 domain is appended to the C terminus of some amphioxus CYP74, such as Bf-EAS (Supplementary Fig. 26). This domain is also located in cAOS–LOX, between the cAOS

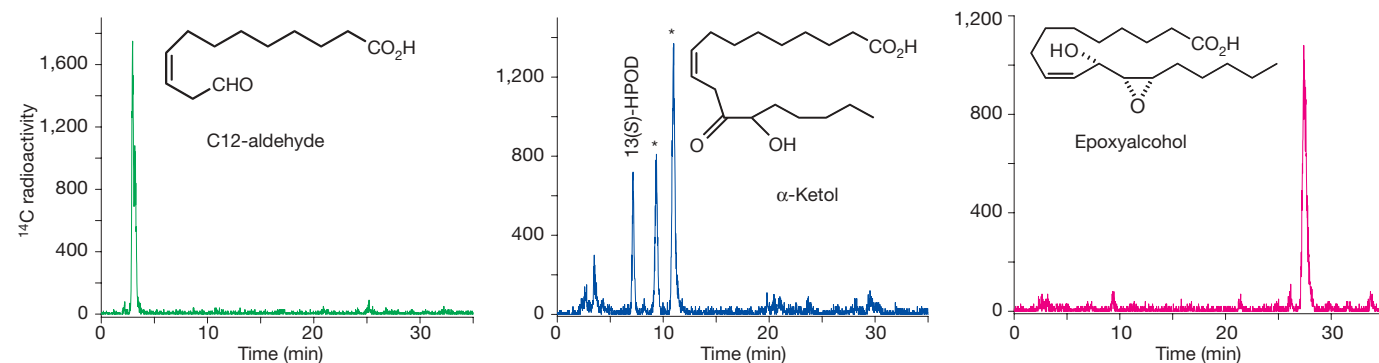


Figure 5 | Discovery of CYP74 in bacteria and animals. Recombinant CYP74 proteins from *M. nodulans* (left panel), *A. palmata* (middle panel) and *B. floridae* (right panel) were incubated with [1-¹⁴C]13(S)-HPOD (200 μ M) at 23 °C for 15 min. After derivatization, radio-HPLC was used to quantify product formation. *M. nodulans* and *A. palmata* enzymes generated

C12-aldehyde and α -ketol, respectively. Peaks marked by an asterisk correspond to the *syn/anti* isomers of α -ketol. *B. floridae* CYP74 catalysed the formation of epoxycarboxaldehyde (12R,13S-epoxy-11S-hydroxy-9Z-octadecenoic acid). The latter is the 11S-hydroxy derivative of (–)-vernolic acid, whose mode of interaction with At-AOS is shown in Fig. 3a.

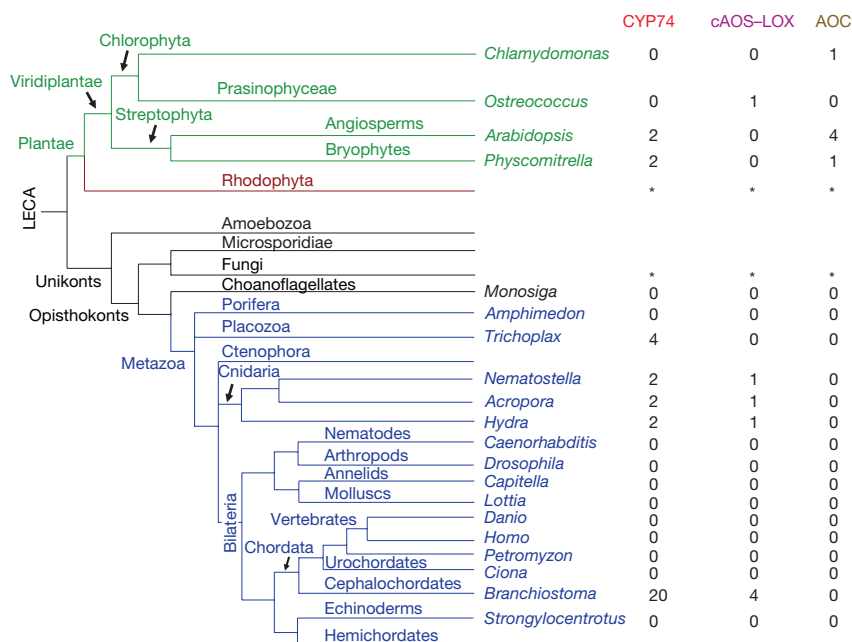


Figure 6 | Common ancestry of oxylipin biosynthetic enzymes. Lineage-specific expansion of CYP74, cAOS-LOX and AOC genes is shown (left panel). CYP74 collectively refers to fatty-acid-hydroperoxide-metabolizing non-monoxygenase cytochrome P450 family members that incorporate both a nine-residue insert in the proximal Cys ligand loop and the active site residues of At-AOS. There are 20 CYP74 genes in amphioxus, but only 4 that encode cAOS-LOX (right panel). A zero indicates that the gene was not found in the specified organism. An asterisk denotes the ability to biosynthesize jasmonic acid, but we have been unable to locate CYP74 or AOC genes in the available rhodophyte and fungal genomes (Supplementary Discussion). LECA, last eukaryotic common ancestor.

and LOX coding regions (see below). Fifth, Mn-HPL is N-terminally fused to a cyclooxygenase-like (COX-like) gene (Supplementary Figs 27–29 and Supplementary Discussion).

We investigated whether CYP74 and cAOS-LOX coexist. We identified both genes in cnidarians and cephalochordates (Fig. 6 and Supplementary Figs 30 and 31). Of the four cAOS-LOX paralogues found in the latter, one resembles the *Anabaena* enzyme (His ligand)⁴¹, another is similar to coral cAOS (Tyr ligand)³⁸ and the remaining two are new. We also obtained evidence to suggest that cAOS-LOX arose from a gene fusion event. In *Plesiocystis pacifica*, cAOS and LOX are encoded by two separate genes that form an operon (Supplementary Fig. 31). Because a long-chain polyunsaturated fatty acid (C20:4) is produced by this marine myxobacterium (Supplementary Discussion), it may serve as the substrate for generating allene oxide. Although we could not find an allene oxide cyclase (AOC) (Supplementary Fig. 1) orthologue in *P. pacifica*, we located it in other myxobacteria, including in *Stigmatella aurantiaca* (Supplementary Figs 32 and 33). However, the latter does not encode CYP74 genes.

Together, our phylogenetic analysis indicates that CYP74 was present in the last common ancestor of plants and animals (Fig. 6). CYP74 genes persist in plants and in basal extant animals, including *Trichoplax*, corals, hydra and sea anemone. Even though they experienced considerable expansion in the most basal chordate (amphioxus), they were lost in vertebrates and their closest invertebrate relatives (tunicates). The propagation of these genes as multiple-copy clusters also suggests that a duplication event may have occurred in the ancestral CYP74 before the divergence of plants and animals (Supplementary Discussion). The coexistence of cAOS-LOX and CYP74 has been preserved in the cnidarian and cephalochordate lineages, but was eliminated during placozoan evolution. This illustrates that both genes were present in an animal before the cnidarian-bilaterian split⁴⁴, which occurred approximately 600 million years ago.

What is the physiological significance of CYP74 activity in bacteria and animals? This is the first time a GLV biosynthetic enzyme has been found in bacteria. Because *Methylobacterium* strains contribute to enhancing the flavour of strawberries⁴⁵, our discovery of Mn-HPL has direct implications for understanding how plant-growth-promoting rhizobacteria^{46,47} mediate their actions. Notably, *M. nodulans* is a root-nodule-forming and nitrogen-fixing symbiont of *Crotalaria* species⁴⁸, the methylotrophic metabolism of which has been shown to promote optimal plant growth⁴⁹. At this time, we

are unable to establish whether bacterial HPL was acquired by means of horizontal gene transfer from plants (Supplementary Discussion). Where animals are concerned, coexistence of cAOS-LOX and AOS in *Acropora* suggests that they are part of different biosynthetic pathways. Also, the catalytic activity of Bf-EAS is reminiscent of epoxyl-cohol formation from arachidonic acid by means of the 12R-LOX-eLOX3 pathway⁵⁰ in mammals. Related to this, we have established that C18 polyunsaturated fatty acids are abundant in amphioxus (Supplementary Table 6). Therefore, substrate availability is not an issue for the animal CYP74 enzymes. More interestingly, we have identified a leukotriene A4 hydrolase orthologue in animals that harbour CYP74 genes (Supplementary Fig. 34 and Supplementary Discussion). This lends additional support for fatty acid epoxide metabolism in marine invertebrates.

This paper explains how unusual cytochrome P450 enzymes convert fatty acid hydroperoxides into a broad spectrum of molecules that have beneficial applications in agriculture and medicine. It provides essential information for genetically modifying GLV biosynthesis to, among other things, enhance crop plants' resistance against biotic and abiotic stresses, fine-tune bitterness intensity in virgin olive oils, alter floral, fruit and vegetable flavours, and dissect plant volatile signalling mechanisms (Supplementary Discussion). In addition, it opens new vistas for understanding how biofertilizers work, as well as for exploring the biological functions of products derived from animal CYP74 enzymes.

METHODS SUMMARY

The At-AOS structure was determined by multiwavelength anomalous dispersion methods using selenomethionine-labelled protein. Difference Fourier techniques were used to solve the structures of the mutants as well as the complexes. Product formation in enzyme-catalysed reactions was assessed by gas-chromatography/mass spectrometry and radio-HPLC.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 November 2007; accepted 1 August 2008.

Published online 20 August 2008.

1. Liechti, R. & Farmer, E. E. Jasmonate biochemical pathway. *Sci. STKE* 2006, cm3 (2006).
2. Liechti, R., Gfeller, A. & Farmer, E. E. Jasmonate signaling pathway. *Sci. STKE* 2006, cm2 (2006).
3. Wasternack, C. Jasmonates: an update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann. Bot. (Lond.)* 100, 681–697 (2007).

4. Flescher, E. Jasmonates in cancer therapy. *Cancer Lett.* **245**, 1–10 (2007).
5. Matsui, K. Green leaf volatiles: hydroperoxide lyase pathway of oxylipin metabolism. *Curr. Opin. Plant Biol.* **9**, 274–280 (2006).
6. Baldwin, I. T., Halitschke, R., Paschold, A., von Dahl, C. C. & Preston, C. A. Volatile signaling in plant–plant interactions: “talking trees” in the genomics era. *Science* **311**, 812–815 (2006).
7. Xie, D. X., Feys, B. F., James, S., Nieto-Rostro, M. & Turner, J. G. COI1: an *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science* **280**, 1091–1094 (1998).
8. Li, L. *et al.* The tomato homolog of CORONATINE-INSENSITIVE1 is required for the maternal control of seed maturation, jasmonate-signaled defense responses, and glandular trichome development. *Plant Cell* **16**, 126–143 (2004).
9. Kessler, A., Halitschke, R. & Baldwin, I. T. Silencing the jasmonate cascade: induced plant defenses and insect populations. *Science* **305**, 665–668 (2004).
10. Gershenzon, J. Plant volatiles carry both public and private messages. *Proc. Natl Acad. Sci. USA* **104**, 5257–5258 (2007).
11. Funk, C. D. Prostaglandins and leukotrienes: advances in eicosanoid biology. *Science* **294**, 1871–1875 (2001).
12. Keirse, M. J. Natural prostaglandins for induction of labor and preinduction cervical ripening. *Clin. Obstet. Gynecol.* **49**, 609–626 (2006).
13. Zimmerman, D. C. A new product of linoleic acid oxidation by a flaxseed enzyme. *Biochem. Biophys. Res. Commun.* **23**, 398–402 (1966).
14. Hamberg, M. Mechanism of corn hydroperoxide isomerase: detection of 12,13(S)-oxido-9(Z),11-octadecadienoic acid. *Biochim. Biophys. Acta* **920**, 76–84 (1987).
15. Brash, A. R., Baertschi, S. W., Ingram, C. D. & Harris, T. M. Isolation and characterization of natural allene oxides: unstable intermediates in the metabolism of lipid hydroperoxides. *Proc. Natl Acad. Sci. USA* **85**, 3382–3386 (1988).
16. Song, W. C. & Brash, A. R. Purification of an allene oxide synthase and identification of the enzyme as a cytochrome P-450. *Science* **253**, 781–784 (1991).
17. Song, W. C., Baertschi, S. W., Boeglin, W. E., Harris, T. M. & Brash, A. R. Formation of epoxycalcohols by a purified allene oxide synthase. Implications for the mechanism of allene oxide synthesis. *J. Biol. Chem.* **268**, 6293–6298 (1993).
18. Tijet, N. & Brash, A. R. Allene oxide synthases and allene oxides. *Prostaglandins Other Lipid Mediat.* **68–69**, 423–431 (2002).
19. Park, J. H. *et al.* A knock-out mutation in allene oxide synthase results in male sterility and defective wound signal transduction in *Arabidopsis* due to a block in jasmonic acid biosynthesis. *Plant J.* **31**, 1–12 (2002).
20. Vick, B. A. & Zimmerman, D. C. Lipoxygenase and hydroperoxide lyase in germinating watermelon seedlings. *Plant Physiol.* **57**, 780–788 (1976).
21. Grechkin, A. N. & Hamberg, M. The “heterolytic hydroperoxide lyase” is an isomerase producing a short-lived fatty acid hemiacetal. *Biochim. Biophys. Acta* **1636**, 47–58 (2004).
22. Itoh, A. & Howe, G. A. Molecular cloning of a divinyl ether synthase. Identification as a CYP74 cytochrome P-450. *J. Biol. Chem.* **276**, 3620–3627 (2001).
23. Grechkin, A. N. Hydroperoxide lyase and divinyl ether synthase. *Prostaglandins Other Lipid Mediat.* **68–69**, 457–470 (2002).
24. Stumpe, M. & Feussner, I. Formation of oxylipins by CYP74 enzymes. *Phytochem. Rev.* **5**, 347–357 (2006).
25. Guengerich, F. P. Cytochrome P450 enzymes in the generation of commercial products. *Nat. Rev. Drug Discov.* **1**, 359–366 (2002).
26. Denisov, I. G., Makris, T. M., Sligar, S. G. & Schlichting, I. Structure and chemistry of cytochrome P450. *Chem. Rev.* **105**, 2253–2277 (2005).
27. Poulos, T. L. & Johnson, E. F. in *Cytochrome P450: Structure, Mechanism and Biochemistry* 3rd edn (ed. P. R. Ortiz de Montellano) 87–114 (Plenum, 2005).
28. Williams, P. A., Cosme, J., Sridhar, V., Johnson, E. F. & McRee, D. E. Mammalian microsomal cytochrome P450 monooxygenase: structural adaptations for membrane binding and functional diversity. *Mol. Cell* **5**, 121–131 (2000).
29. Froehlich, J. E., Itoh, A. & Howe, G. A. Tomato allene oxide synthase and fatty acid hydroperoxide lyase, two cytochrome P450s involved in oxylipin metabolism, are targeted to different membranes of chloroplast envelope. *Plant Physiol.* **125**, 306–317 (2001).
30. Vidi, P. A. *et al.* Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. *J. Biol. Chem.* **281**, 11225–11234 (2006).
31. Joo, H., Lin, Z. & Arnold, F. H. Laboratory evolution of peroxide-mediated cytochrome P450 hydroxylation. *Nature* **399**, 670–673 (1999).
32. Bell, C. D., Soltis, D. E. & Soltis, P. S. The age of the angiosperms: a molecular timescale without a clock. *Evolution Int. J. Org. Evolution* **59**, 1245–1258 (2005).
33. White, R. E. & Coon, M. J. Oxygen activation by cytochrome P-450. *Annu. Rev. Biochem.* **49**, 315–356 (1980).
34. Smith, D. M., Nicolaidis, A., Golding, B. T. & Radom, L. Ring opening of the cyclopropylcarbonyl radical and its N- and O-substituted analogs: A theoretical examination of very fast unimolecular reactions. *J. Am. Chem. Soc.* **120**, 10223–10233 (1998).
35. Groves, J. T. The bioinorganic chemistry of iron in oxygenases and supramolecular assemblies. *Proc. Natl Acad. Sci. USA* **100**, 3569–3574 (2003).
36. Green, M. T., Dawson, J. H. & Gray, H. B. Oxoiron(IV) in chloroperoxidase compound II is basic: implications for P450 chemistry. *Science* **304**, 1653–1656 (2004).
37. Ogliaro, F., de Visser, S. P. & Shaik, S. The ‘push’ effect of the thiolate ligand in cytochrome P450: a theoretical gauging. *J. Inorg. Biochem.* **91**, 554–567 (2002).
38. Koljak, R., Boutaud, O., Shieh, B. H., Samel, N. & Brash, A. R. Identification of a naturally occurring peroxidase–lipoxygenase fusion protein. *Science* **277**, 1994–1996 (1997).
39. Oldham, M. L., Brash, A. R. & Newcomer, M. E. The structure of coral allene oxide synthase reveals a catalase adapted for metabolism of a fatty acid hydroperoxide. *Proc. Natl Acad. Sci. USA* **102**, 297–302 (2005).
40. Lohelaid, H. *et al.* Identification of a functional allene oxide synthase–lipoxygenase fusion protein in the soft coral *Gersemia fruticosa* suggests the generality of this pathway in octocorals. *Biochim. Biophys. Acta* **1780**, 315–321 (2008).
41. Schneider, C. *et al.* Enzymatic synthesis of a bicyclobutane fatty acid by a hemoprotein lipoxygenase fusion protein from the cyanobacterium *Anabaena* PCC 7120. *Proc. Natl Acad. Sci. USA* **104**, 18941–18945 (2007).
42. Nelson, D. R. Plant cytochrome P450s from moss to poplar. *Phytochem. Rev.* **5**, 193–204 (2006).
43. Rensing, S. A. *et al.* The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64–69 (2008).
44. Cartwright, P. & Collins, A. Fossils and phylogenies: integrating multiple lines of evidence to investigate the origin of early metazoan lineages. *Integr. Comp. Biol.* **47**, 744–751 (2007).
45. Zabetakis, I. Enhancement of flavour biosynthesis from strawberry (*Fragaria ananassa*) callus cultures by *Methylobacterium* species. *Plant Cell Tissue Organ Cult.* **50**, 179–183 (1997).
46. Van Loon, L. C. & Bakker, P. A. in *Plant-associated Bacteria* (ed. Gnanamanickam, S. S.) 269–316 (Springer, 2006).
47. Beckers, G. J. & Conrath, U. Priming for stress resistance: from the lab to the field. *Curr. Opin. Plant Biol.* **10**, 425–431 (2007).
48. Renier, A. *et al.* Symbiotic properties of *Methylobacterium nodulans* ORS 2060T: A classic process for an atypical symbiont. *Soil Biol. Biochem.* **40**, 1404–1412 (2008).
49. Jourand, P. *et al.* Role of methylothrophy during symbiosis between *Methylobacterium nodulans* and *Crotalaria podocarpa*. *Mol. Plant Microbe Interact.* **18**, 1061–1068 (2005).
50. Brash, A. R., Yu, Z., Boeglin, W. E. & Schneider, C. The hepxoilin connection in the epidermis. *FEBS J.* **274**, 3494–3502 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements C.S.R. is grateful to the late N. Natarajaratnam and the late T. Traylor for encouragement; R. Cudney for nonanoyl-N-hydroxyethylglucamide (HEGA-9); K. Matsui for tomato HPL complementary DNA; K. Back and D. Park for *Arabidopsis* and rice cDNA libraries; L. Holland and J. Langeland for the amphioxus cDNA library; M. Medina for *A. palmata* expressed sequence tags; C. Marx and L. Moulin for *M. nodulans* cells; R. Müller for *S. aurantiaca* genomic DNA; L. Roman and B. Masters for the pCW_{ori}⁺ vector; J. Navarro for the single crystal microspectrophotometer; T. Doukov, S. Soltis, A. Cohen and J. Charles for help with acquiring electronic absorption spectra of single crystals; S. Veeraraghavan for generating Fig. 5; R. Bach, T. Bach, A. Beckwith, W. Bernhard, D. Curran, A. Davies, T. Dibble, J. Finnerty, D. Fleischman, J. Froehlich, J. Groves, L. Holland, P. Holland, J. Howieson, H. Kaplan, D. Nelson, M. Newcomb, P. Ortiz de Montellano, N. Porter, T. Poulos, M. Sibi, S. Veeraraghavan and D. Whalen for discussions; Joint Genome Institute for access to sequence data; Stanford Synchrotron Radiation Laboratories (beam lines 9–2 and 11–1, T. Doukov and L. Dunn) and the Advanced Light Source (beam line 8.3.1, J. Holton, G. Meigs and J. Tanamachi; beam lines 8.2.1 and 8.2.2, C. Ralston) for beam time and assistance. This work is supported by Pew Charitable Trusts through a Pew Scholar Award (C.S.R.), The Robert A. Welch Foundation (C.S.R.), The National Institutes of Health (C.S.R.), a Beginning Grant in Aid from the American Heart Association (D.-S.L.), and an INSERM Avenir Grant sponsored by La Fondation pour la Recherche Médicale (P.N.).

Author Contributions C.S.R. designed the research. D.-S.L. overexpressed and purified all the proteins used in this work; D.-S.L. and P.N. measured enzyme kinetic data; D.-S.L. generated the crystals; D.-S.L. and C.S.R. collected X-ray diffraction data; P.N. and C.S.R. solved the structures; P.N. did structure refinements; C.S.R. carried out bioinformatic and phylogenetic analyses; M.H. performed GC–MS and radio-HPLC measurements, and determined the structures of the reaction products; and C.S.R. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Coordinates and structure factors have been deposited in the RCSB Protein Data Bank under the following accession codes: 3CLI (At-AOS), 3DSI (At-AOS in complex with 13-HOT), 2RCH (At-AOS in complex with 13-HOD), 2RCL (At-AOS in complex with 12R,13S-vernolic acid), 2RCM (At-AOS(F137L)), 3DSJ (At-AOS(F137L) in complex with 13-HOD), and 3DSK (At-AOS(F137L) in complex with 12R,13S-vernolic acid). Nucleotide sequences of Ap-AOS, Bf-EAS and Mn-HPL have been deposited in Genbank under accession numbers EU541487, EU555186 and EU887514, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.S.R. (ramancs@gmail.com).

METHODS

Protein expression and purification. We used the Johnson–Waterman method^{51–54} to overproduce all the cytochromes P450 and their variants described in this work. AOS gene was PCR-amplified using *A. thaliana* and *O. sativa* cDNA libraries (gift from K. W. Back and D. S. Park) as templates. *Lycopersicon esculentum* (tomato) HPL cDNA was a gift from K. Matsui. For overexpression of At-AOS (Genbank accession number: Q96242) in *E. coli*, we generated an insert that lacks the first 32 residues, encompassing the chloroplast targeting peptide⁵⁵. We also added seven residues (MAKKTSS) to the N terminus for improving protein solubility and a C-terminal tetra-histidine tag to facilitate purification. *O. sativa* (rice) AOS (Genbank accession number: AAL17675) and *L. esculentum* HPL (Genbank accession number: AAF67142) amplicons were generated in the same manner, but with the exception that only 21 N-terminal residues were removed. Bacterial and animal CYP74 orthologues were overproduced with intact N-termini. *Methylobacterium nodulans* HPL was PCR amplified from bacterial cells provided by C. Marx and L. Moulin. *Branchiostoma floridae* EAS was PCR amplified using cDNA library⁵⁶ (gift from L. Holland and J. Langeland) as the template. This insert includes the C-terminal LH2 domain. An expressed sequence tag (GENBANK accession number: DR983439; EST name: JGI_AOKF1116.fwd; gift from M. Medina) was used to clone *Acropora palmata* AOS. All inserts were ligated into the pCW_{ori} vector (gift of L. Roman and B. S. Masters), sequence verified (Lone Star Labs), and transformed into *Escherichia coli* BL21 (DE3) cells. Post induction, the cells were harvested and the lysate was used in subsequent purification steps. The choice of detergent was empirically determined by means of screening and optimization. We adapted the Johnson method⁵⁴ to purify the P450s in three steps: (1) affinity chromatography (Ni-NTA, Qiagen) in the presence of 0.6–1% *n*-octyl- β -D-glucopyranoside (Anatrace); (2) ion exchange chromatography; and (3) gel filtration (Superdex 200 16/60; GE Healthcare). Haem content was assessed by means of pyridine hemochrome measurements⁵⁷. Labelled protein was purified in a similar manner except that the cells were grown in minimal media supplemented with selenomethionine. Before crystallization At-AOS was detergent exchanged into nonanoyl-*N*-hydroxyethylglucamide (HEGA-9).

S. aurantiaca allene oxide cyclase (AOC) was PCR amplified from genomic DNA (gift from R. Müller). For overproduction and purification, we adapted methods described for *Arabidopsis* AOC⁵⁸.

Activity measurements. AOS, HPL and their variants (0.2–4 μ M) were incubated with hydroperoxides (200 μ M) at 23 °C for 15 min. Part of the reaction product was treated with 3 vol of 30 mM *O*-methylhydroxylamine in methanol at 23 °C for 2 h to generate *O*-methyloxime derivatives of short-chain aldehydes and other carbonyl-containing oxylipins. Additional derivatization of products extracted with diethyl ether was performed by consecutive treatments with ethereal diazomethane and a 2:1:2 (vol/vol/vol) mixture of trimethylchlorosilane, hexamethyldisilazane and pyridine to generate methyl esters from carboxylic acids and trimethylsilyl ethers from alcohols, respectively. As a control, hydroperoxides were also incubated with preparations of native maize AOS⁵⁹, garlic DES⁶⁰ and guava HPL⁶¹.

Oxylipin profiles generated from the incubation of hydroperoxides with CYP74 enzymes were determined by GC–MS analysis run in the scan mode (*m/z* 50–600) and using the authentic compounds as references (Lipidox Co.). The selected ion monitoring (sim) mode was used for sensitive and specific detection. For analysis of oxylipins the following ions and compounds were used. 13-HPOD-derived products: *m/z* 412, 13-hydroxy-12-oxo-9(*Z*)-octadecenoic acid (AOS); *m/z* 308, etheroleic acid (DES); *m/z* 224, 12-oxo-9(*Z*)-dodecenoic acid (HPL). 9-HPOD-derived products: *m/z* 412, 9-hydroxy-10-oxo-12(*Z*)-octadecenoic acid (AOS); *m/z* 308, colneleic acid (DES); *m/z* 184 and 152, 9-oxononanoic acid (HPL). 13-HPOT-derived products: *m/z* 410, 13-hydroxy-12-oxo-9(*Z*),15(*Z*)-octadecadienoic acid (AOS); *m/z* 306, 12-oxophytodienoic acid (AOS); *m/z* 306, etherolenic acid (DES); *m/z* 224, 12-oxo-9(*Z*)-dodecenoic acid (HPL). 11-HPHT-derived products: *m/z* 382, 11-hydroxy-10-oxo-7(*Z*),13(*Z*)-hexadecadienoic acid; *m/z* 278, dinor-12-oxophytodienoic acid (AOS); *m/z* 278, dinor-etherolenic acid (DES); and *m/z* 196, 10-oxo-7(*Z*)-decanoic acid (HPL). The abundance of products generated in the various incubations was estimated by integration of the intensities of mass spectral ions.

Oxylipin production by recombinant CYP74 enzymes from *M. nodulans*, *A. palmata* and *B. floridae* was measured by radio-HPLC methods. Initially, 0.2–4 μ M of the enzyme was incubated with [¹⁻¹⁴C]13(*S*)-HPOD (200 μ M) at 23 °C for 15 min. This was followed by derivatization of carbonyl groups with 30 mM *O*-methyl hydroxylamine in methanol at room temperature (23 °C) for 2 h. After solvent extraction and methyl-esterification, normal phase HPLC was performed using 0.6% 2-propanol-hexane (0–15 min) and 1.2% 2-propanol-hexane (15–35 min) for elution.

Enzyme kinetics. Hydroperoxide consumption by AOS and its mutants was quantified using ultraviolet spectrophotometry⁶², by following the decrease in

absorption at 235 nm stemming from the loss of conjugated diene system of the substrate. Reactions were performed at 23 °C in 0.5 ml of 50 mM sodium phosphate buffer, pH 7. Initial velocity measurements used fixed enzyme concentration (wild-type AOS, 6 nM; Phe to Leu mutants, 1.5 nM; and At-AOS(N321Q), 30 nM) while varying the 13(*S*)-HPOD concentration. Activity slopes were obtained under initial velocity conditions, in which less than 10% of the substrate has been converted into product.

Crystallization and data collection. Crystals were obtained by the sitting-drop vapour diffusion method at 20 °C. At-AOS crystals were grown from 100 mM Tris-HCl, pH 7.5, 15–20% polyethylene glycol (PEG) 3350, and 39 mM HEGA-9. Substrate and intermediate analogues were premixed with the protein solution before generating crystals of the complexes. Streak seeding was used to grow crystals of AOS variants.

Because X-rays are capable of reducing the haem iron during data collection⁶³, we have used three distinct strategies for minimizing X-ray damage to crystals: (1) use a data collection strategy that generates a complete diffraction data set from a single crystal with the least amount of X-ray exposure (for example, in the At-AOS case we were able to collect the entire data set within 2 min with each frame being exposed for only 0.5 s; total dose received = 3.8×10^5 grey or 2% of Henderson limit); (2) perform data collection at a shorter wavelength (below 0.83 Å)⁶⁴; (3) use single crystals that were co-crystallized in the presence of 10 mM potassium ferricyanide⁶⁵. Visible absorption spectra of At-AOS crystals were obtained using a microspectrophotometer setup (4DXray Systems) adapted for use at the Stanford Synchrotron Radiation Laboratories.

Diffraction data were collected (100 K) at SSRL BL 9–2, BL 11–1, and ALS BL 8.3.1, BL 8.2.1 and BL 8.2.2. Glycerol was used as the cryoprotectant. Single wavelength anomalous diffraction and multiwavelength anomalous diffraction data sets were collected at the iron and selenium edges (Supplementary Tables 1 and 2). Data were integrated and scaled with MOSFLM⁶⁶/SCALA⁶⁷ or HKL2000 (ref. 68).

Structure determination and refinement. Se-MAD data was used to obtain phase information. SHELXS⁶⁹ was used to locate the anomalous scatterers (16 Se). Refinement of heavy-atom parameters and phase calculation were done using SHARP⁷⁰. Solvent flattening with SOLOMON⁷¹ and phase extension yielded electron density maps into which most residues were built with the program O⁷². The model was further refined with REFMAC5 (ref. 73) using a maximum-likelihood target function and translation/libration/screw refinement⁷⁴. Subsequent model building allowed us to generate the complete model (Supplementary Tables 1 and 2). Structures of ligand-bound complexes were obtained by means of difference Fourier methods. Figures were generated using PYMOL⁷⁵.

51. Barnes, H. J., Arlotto, M. P. & Waterman, M. R. Expression and enzymatic activity of recombinant cytochrome P450 17 α -hydroxylase in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **88**, 5597–5601 (1991).
52. Richardson, T. H. et al. Purification and characterization of recombinant-expressed cytochrome P450 2C3 from *Escherichia coli*: 2C3 encodes the 6 β -hydroxylase deficient form of P450 3b. *Arch. Biochem. Biophys.* **300**, 510–516 (1993).
53. von Wachenfeldt, C., Richardson, T. H., Cosme, J. & Johnson, E. F. Microsomal P450 2C3 is expressed as a soluble dimer in *Escherichia coli* following modification of its N-terminus. *Arch. Biochem. Biophys.* **339**, 107–114 (1997).
54. Wester, M. R., Stout, C. D. & Johnson, E. F. Purification and crystallization of N-terminally truncated forms of microsomal cytochrome P450 2C5. *Methods Enzymol.* **357**, 73–79 (2002).
55. Hughes, R. K. et al. Allene oxide synthase from *Arabidopsis thaliana* (CYP74A1) exhibits dual specificity that is regulated by monomer–micelle association. *FEBS Lett.* **580**, 4188–4194 (2006).
56. Langeland, J. A., Tomsa, J. M., Jackman, J. W. R. & Kimmel, C. B. An amphioxus snail gene: expression in paraxial mesoderm and neural plate suggests a conserved role in patterning the chordate embryo. *Dev. Genes Evol.* **208**, 569–577 (1998).
57. Berry, E. A. & Trumpower, B. L. Simultaneous determination of hemes a, b, and c from pyridine hemochrome spectra. *Anal. Biochem.* **161**, 1–15 (1987).
58. Hofmann, E., Zerbe, P. & Schaller, F. The crystal structure of *Arabidopsis thaliana* allene oxide cyclase: insights into the oxylipin cyclization reaction. *Plant Cell* **18**, 3201–3217 (2006).
59. Hamberg, M. & Fahlstadius, P. Allene oxide cyclase: a new enzyme in plant lipid metabolism. *Arch. Biochem. Biophys.* **276**, 518–526 (1990).
60. Grechkin, A. N., Fazliev, F. N. & Mukhtarova, L. S. The lipoygenase pathway in garlic (*Allium sativum* L.) bulbs: detection of the novel divinyl ether oxylipins. *FEBS Lett.* **371**, 159–162 (1995).
61. Grechkin, A. N. & Hamberg, M. The “heterolytic hydroperoxide lyase” is an isomerase producing a short-lived fatty acid hemiacetal. *Biochim. Biophys. Acta* **1636**, 47–58 (2004).
62. Brash, A. R. & Song, W. C. Detection, assay, and isolation of allene oxide synthase. *Methods Enzymol.* **272**, 250–259 (1996).

63. Beitlich, T., Kuhnel, K., Schulze-Bries, C., Shoeman, R. L. & Schlichting, I. Cryoradiolytic reduction of crystalline heme proteins: analysis by UV-Vis spectroscopy and X-ray crystallography. *J. Synchrotron Radiat.* **14**, 11–23 (2007).
64. Schlichting, I. Crystallographic structure determination of unstable species. *Acc. Chem. Res.* **33**, 532–538 (2000).
65. Lancaster, C. R., Kroger, A., Auer, M. & Michel, H. Structure of fumarate reductase from *Wolinella succinogenes* at 2.2 Å resolution. *Nature* **402**, 377–385 (1999).
66. Leslie, A. G. W. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26. (Warrington, UK, Daresbury Lab, 1992).
67. Collaborative Computational Project Number 4. The CCP4 suite programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
68. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–325 (1997).
69. Sheldrick, G. M. Patterson superpositioning and ab initio phasing. *Methods Enzymol.* **276**, 628–641 (1997).
70. de La Fortelle, E. & Bricogne, G. Maximum-likelihood heavy atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
71. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D* **52**, 30–42 (1996).
72. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models into electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
73. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
74. Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* **57**, 122–133 (2001).
75. DeLano, W. L. The PyMOL Molecular Graphics System <<http://www.pymol.org>> (DeLano Scientific, 2002).

A periodicity of ~ 1 hour in X-ray emission from the active galaxy RE J1034+396

Marek Gierliński¹, Matthew Middleton¹, Martin Ward¹ & Chris Done¹

Active galactic nuclei and quasars are thought to be scaled-up versions of Galactic black hole binaries, powered by accretion onto supermassive black holes with masses of 10^6 – $10^9 M_\odot$, as opposed to the $\sim 10 M_\odot$ in binaries (here M_\odot is the solar mass). One example of the similarities between these two types of systems is the characteristic rapid X-ray variability seen from the accretion flow¹. The power spectrum of this variability in black hole binaries consists of a broad noise with multiple quasi-periodic oscillations superimposed on it. Although the broad noise component has been observed in many active galactic nuclei^{2,3}, there have hitherto been no significant detections of quasi-periodic oscillations^{4–6}. Here we report the discovery of an ~ 1 -hour X-ray periodicity in a bright active galaxy, RE J1034+396. The signal is highly statistically significant (at the 5.6σ level) and very coherent, with quality factor $Q > 16$. The X-ray modulation arises from the direct vicinity of the black hole.

RE J1034+396 (also known as KUG 1031+398) is a nearby (redshift $z = 0.042$) active galaxy, spectroscopically classified as a narrow-line Seyfert 1 (NLS1). These objects have strong emission lines produced by high-density gas ionized by the ultraviolet and X-ray radiation from the accretion flow. These lines are rather narrow compared to the velocity widths seen in more typical broad-line active galactic nuclei (AGN). This fact, together with other evidence, has led to the suggestion that they host supermassive black holes that are less massive than those inferred in a typical AGN of similar luminosity⁷.

From a long (91 ks) observation using the X-ray satellite XMM-Newton, we extracted a light curve for RE J1034+396 (Fig. 1) over the energy band 0.3–10 keV. Even by eye it shows an evident periodic oscillation. To test more rigorously for the presence of a periodic signal, we folded the light curve with various trial periods and analysed the root-mean-square (r.m.s.) amplitude of the resulting pulse profile as a function of the period. We found a strong peak at $3,730 \pm 130$ s (full-width at half-maximum). We used the best-fitting period to plot the expected times of minima in Fig. 1 (dotted vertical lines). This shows that the periodicity changes its character at around $t_0 = 25$ ks. After that time, the troughs in the light curve follow the predicted minima very well, for almost 16 cycles, indicating a highly coherent signal, but before t_0 the troughs are shifted in phase and there are occasional additional minima. This shows that the feature is not a true periodicity, but that it wanders in phase, amplitude and/or frequency, as seen in the quasi-periodic oscillations (QPOs) in black hole binaries (BHBs)⁸. Hence, we will refer to this signal as a QPO.

We concentrate first on the coherent part of the light curve (segment 2 in Fig. 1). Figure 2 shows this light curve segment folded with the best-fitting period, and Fig. 3 shows its periodogram with a strong peak at $\sim 2.7 \times 10^{-4}$ Hz. In order to quantify the statistical significance of the peak, we adopt the method proposed⁹ to test the significance of periodicities against an underlying continuum. This method

involves dividing the periodogram by the best-fitting power law and using the known distribution of the periodogram ordinates to estimate the likelihood of observing a given peak. The confidence limits (3σ and 99.99%) shown in Fig. 3 are calculated including the uncertainties in the power-law model. The QPO is well above these limits, with a chance probability of a spurious signal of 2×10^{-8} (corresponding to the $\sim 5.6\sigma$ significance level). Even in the total light curve, including segment 1, which has less obvious periodicity, the signal is still significant at the $\sim 3.4\sigma$ level.

This method assumes that the underlying continuum has a power-law shape, but there can be breaks in this continuum, changing the derived significance of the QPO. We tested this with Monte Carlo simulations, generating a series of light curves following a given power spectral distribution¹⁰. The simulated light curves had the same number of bins, mean count rate and variance as the observed light curve. We then calculated periodograms for each of them, and found the power corresponding to the upper 3σ limit in each frequency channel (the maximum significance in this method is 3.8σ

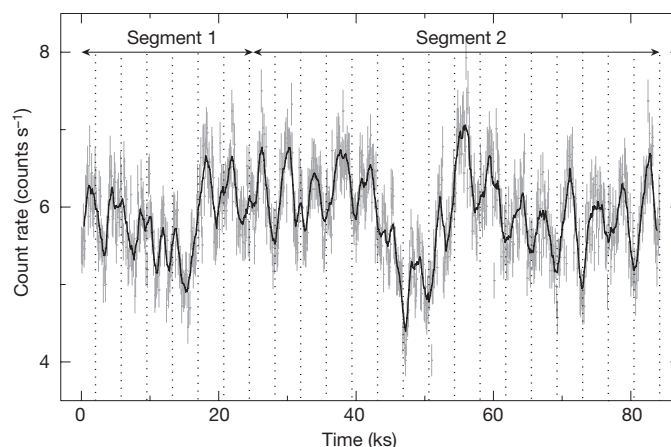


Figure 1 | XMM-Newton light curve of RE J1034+396. The start time of this observation was 2007 May 31, 20:10:12 UTC. We extracted source and background light curves from the PN, MOS1 and MOS2 cameras in the 0.3–10 keV energy band, using 45-arcsec circular selection regions and rejecting the final ~ 7 ks owing to background flares. We then combined the data from all cameras. This gave a mean count rate of 5.9 ± 0.6 and 0.04 ± 0.08 (s.d.) counts s^{-1} for the net and background light curves, respectively. The data points, binned in 100-s intervals, are plotted in grey; error bars, ± 1 s.d. The thick black curve represents the running average over 9 bins around a given bin and shows clear periodicity. The dotted vertical lines show the expected times of minima obtained from folding segment 2 with the period of 3,733 s. In this Letter we mainly analyse segment 2, which shows a periodicity with high coherence. It contains 593 contiguous 100 s bins, and almost 16 full cycles of the periodic signal. The fractional r.m.s. variability (in terms of excess variance²¹) in this segment is $9.2 \pm 0.2\%$.

¹Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK.

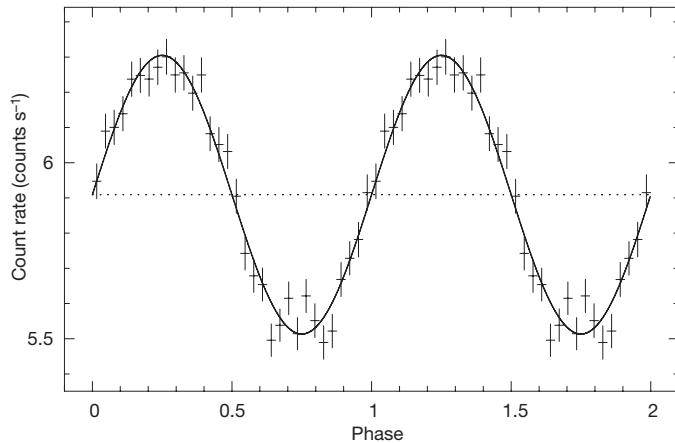


Figure 2 | Folded light curve. We folded the part of the light curve with significant periodicity (segment 2 in Fig. 1) with the period of 3,733 s. Errors are propagated from the unfolded light curve, and represent 1 s.d. Two cycles are plotted for clarity. The solid line represents the best-fitting sinusoid, the dotted line the mean count rate. The amplitude of the sinusoid is $\sim 6.7\%$ of the mean, which corresponds to $\sim 4.7\%$ of the fractional r.m.s. variability in the pulse profile.

owing to the limited number of simulated light curves). The results for a single power-law distribution with the index taken from the best fit to our data (with index uncertainties taken into account) are comparable to the analytical limits shown in Fig. 3, as expected. A broken power law with indices -1 and -2 below and above the break frequency of 2.7×10^{-4} Hz, respectively, decreases the confidence limits around the break, but the QPO remains highly significant ($>3.8\sigma$). Therefore, we conclude that the observed signal at $\sim 2.7 \times 10^{-4}$ Hz in RE J1034+396 is significant at a very high statistical level, irrespective of the assumed model for the continuum power.

The QPO lies within just one frequency bin of the periodogram, which makes it highly coherent, with a quality factor $Q = f/\Delta f > 16$, where f and Δf are the bin frequency and width, respectively. The r.m.s. fractional variability in the QPO is $\sim 4.7\%$ in the 0.3–10 keV energy band, which constitutes about half of the r.m.s. variability in the light curve. The strength of the QPO depends significantly on energy, increasing from $\sim 2\%$ in the 0.2–0.3 keV band to $\sim 10\%$ above 1 keV. There is a time lag of ~ 260 s in the QPO phase between the 2–10 and 0.3–0.4 keV energy bands (softer X-rays lagging behind harder X-rays).

This is the first time that such a strong QPO has been convincingly found in the X-ray light curve of any AGN. Earlier claims all failed a more stringent statistical analysis^{4–6}, such as that used here. The only exception was that claimed for NGC 6814, which turned out to be from a Galactic cataclysmic variable along the line of sight¹¹. However, XMM-Newton is an imaging instrument and has a positional accuracy of ~ 2 –3 arcsec, so considering that the number of X-ray-detected cataclysmic variables is ~ 100 (ref. 12), the probability of a chance superposition with the AGN is vanishingly small ($<10^{-8}$). We therefore conclude that the X-ray source is associated with the galaxy. Furthermore, its X-ray luminosity (4×10^{43} erg s $^{-1}$ in the 0.3–10 keV band) is then too bright for an ultra-luminous binary X-ray source, which have luminosities below $\sim 10^{41}$ erg s $^{-1}$. Therefore it must originate from the AGN.

The black hole mass in RE J1034+396 is not well determined, and different methods of measurement give conflicting results¹³. The virial mass derived from the H β emission line velocity dispersion in the broad-line region¹⁴ is $6.3 \times 10^5 M_{\odot}$, and differs significantly from an estimate of $3.6 \times 10^7 M_{\odot}$ obtained using [O III] as a proxy for the stellar velocity dispersion in the bulge¹⁵. On the other hand, if the mass–velocity dispersion relation in NLS1s¹⁶ differs from that in broad-line galaxies¹⁴, then the latter mass measurement may be

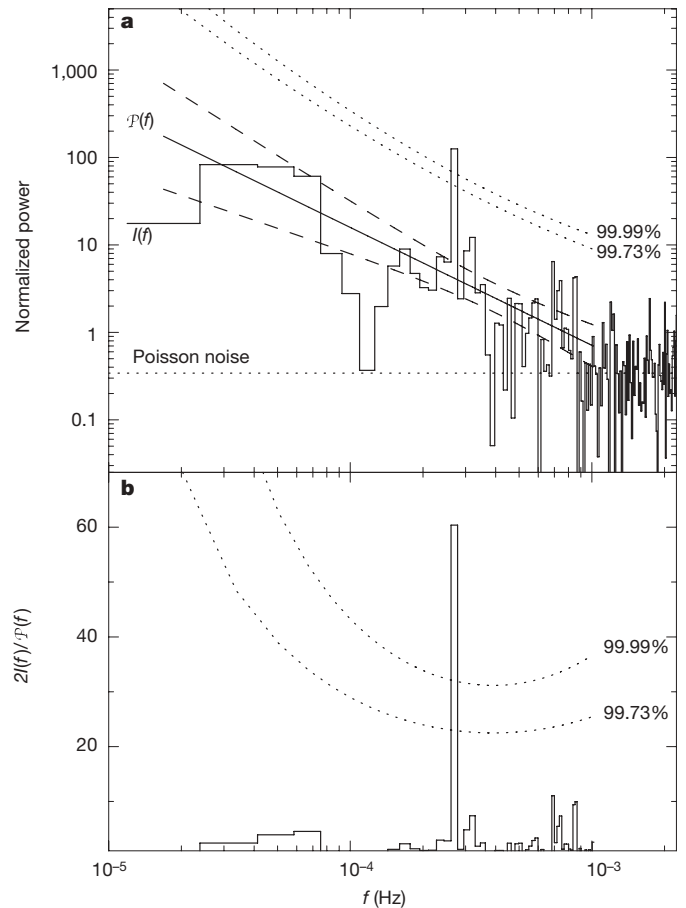


Figure 3 | Power spectral distribution. **a**, Power spectrum (296 data bins), $I(f)$, where I is the power and f is the frequency, normalized to $(\text{r.m.s.}/\text{mean})^2$ per Hz, calculated from segment 2 of the light curve (Fig. 1). The solid line is the de-biased best-fitting (least squares method) power law, $P(f)$, with index -1.35 ± 0.18 . The fit excluded the data above 10^{-3} Hz, dominated by the white noise. We checked that a particular choice of the cut-off did not affect the significance of the peak. The dashed curves represent the uncertainty in the power-law model. The dotted horizontal line shows the expected level of the Poissonian noise. If $P(f)$ represents the true power distribution (which is our null hypothesis), then the quantity $2I(f)/P(f)$ is scattered with a χ^2 distribution with two degrees of freedom^{9,22}. The Kolmogorov–Smirnov test returned the P value of $\sim 84\%$, so the null hypothesis is not rejected at the significance level of, for example, 5%. This shows that a single power law is a good description of the underlying noise process. We also confirm this by Monte Carlo simulations, which allow us to estimate 1 s.d. errors in each frequency channel and calculate $\chi^2_{\nu} = 68.4$ at 58 degrees of freedom. We use the χ^2 distribution of $2I(f)/P(f)$ to calculate the confidence limits on the suspected periodic signal. **b**, Plot of $2I(f)/P(f)$ together with 3σ (99.73%) and 99.99% confidence limits. The same confidence limits are also plotted in **a**. We see that the periodic signal at $\sim 2.7 \times 10^{-4}$ Hz is very strong, significant at a level in excess of 99.99% (the actual significance level is $\sim 5.6\sigma$). The significances are global, that is, corrected for the number of frequencies tested. This result is confirmed by Monte Carlo simulations. The periodic signal is also highly significant in the individual light curves extracted from each separate X-ray camera.

overestimated. Our observed periodicity, if it is related to the Keplerian period of the innermost circular stable orbit, would correspond to a central mass of between 8×10^6 and $9 \times 10^7 M_{\odot}$, for a non-rotating and maximally rotating black hole, respectively.

Galactic BHBs show a huge variety of QPOs, differing in frequency, power and coherence¹⁷. Generally, they can be divided into the low- and high-frequency QPOs, with frequencies <50 and >100 Hz, respectively. The strongest and most coherent low-frequency QPOs are typically seen at frequencies <10 Hz, which, when scaled to the frequency of 2.7×10^{-4} Hz reported here, would imply a black hole mass of less than $\sim 4 \times 10^5 M_{\odot}$ in RE J1034+396. Given

the observed bolometric luminosity (dominated by the soft X-ray/far ultraviolet component) of $\sim 5 \times 10^{44} \text{ erg s}^{-1}$, such a low mass requires that the system be radiating at about 10 times the Eddington limit (L_{Edd}).

High-frequency QPOs are occasionally seen in several BHBs with high coherence⁸. They sometimes occur in pairs, with a frequency ratio of 3:2 (ref. 8). These frequencies appear to be stable, and are regarded as a signature of strong gravity in the vicinity of a rotating black hole¹⁸. A tentative frequency–mass relation, $f_0 = 931 (M/M_\odot)^{-1} \text{ Hz}$, can be derived from three objects⁸. Here f_0 is the fundamental frequency of the pair; that is, the observed frequencies are $2f_0$ and $3f_0$ (the fundamental is not seen). This relation yields the black hole mass in RE J1034+396 of 6.9×10^6 or $1.0 \times 10^7 M_\odot$, depending on whether the observed periodicity corresponds to $2f_0$ or $3f_0$, respectively. This would imply a luminosity of 0.5 or 0.3 L_{Edd} , respectively. All this clearly shows that better mass estimates (using, for example, reverberation mapping or accurate stellar velocity dispersion) are required for RE J1034+396 before the QPO type can be uniquely identified.

Finally we should ask why RE J1034+396 is unique in showing the first convincing evidence for a QPO, given that many AGN have comparable quality X-ray data and have been monitored for similar or longer timescales. Perhaps we have been exceptionally lucky in detecting a QPO with a small duty cycle (the high-frequency QPO in BHBs is seen only occasionally). Alternatively, it may be connected to the fact that RE J1034+396 is extreme even among NLS1 galaxies, with an unusual spectral energy distribution peaking in the far ultraviolet^{19,20}. This component extends into the soft X-ray band pass, but not to the harder X-rays where the QPO is seen. Thus the far ultraviolet/soft X-ray component cannot be directly responsible for the QPO signal, but it may indicate that RE J1034+396 has an extreme mass accretion rate and that this drives both the unusual spectrum and the QPO.

QPOs remain enigmatic, but they clearly contain information about the dynamics of the infalling material. The larger mass of an AGN means that we see fewer cycles of a QPO, but with much higher time resolution compared to BHBs. Therefore, future studies of such phenomena in AGN will shed new light on the origin of QPOs, and in turn, on the behaviour of accretion flows around black holes.

Received 2 June; accepted 17 July 2008.

- McHardy, I. M., Koerding, E., Knigge, C., Uttley, P. & Fender, R. P. Active galactic nuclei as scaled-up Galactic black holes. *Nature* **444**, 730–732 (2006).

- Edelson, R. & Nandra, K. A cutoff in the X-ray fluctuation power density spectrum of the Seyfert 1 galaxy NGC 3516. *Astrophys. J.* **514**, 682–690 (1999).
- Markowitz, A. et al. X-ray fluctuation power spectral densities of Seyfert 1 galaxies. *Astrophys. J.* **593**, 96–114 (2003).
- Benlloch, S., Wilms, J., Edelson, R., Yaqoob, T. & Staubert, R. Quasi-periodic oscillation in Seyfert galaxies: Significance levels. The case of Markarian 766. *Astrophys. J.* **562**, L121–L124 (2001).
- Vaughan, S. & Uttley, P. Where are the X-ray quasi-periodic oscillations in active galaxies? *Mon. Not. R. Astron. Soc.* **362**, 235–244 (2005).
- Vaughan, S. & Uttley, P. Detecting X-ray QPOs in active galaxies. *Adv. Space Res.* **38**, 1405–1408 (2006).
- Boller, T., Brandt, W. N. & Fink, H. Soft X-ray properties of narrow-line Seyfert 1 galaxies. *Astron. Astrophys.* **305**, 53–73 (1996).
- Remillard, R. A. & McClintock, J. E. X-ray properties of black-hole binaries. *Annu. Rev. Astron. Astrophys.* **44**, 49–92 (2006).
- Vaughan, S. A simple test for periodic signals in red noise. *Astron. Astrophys.* **431**, 391–403 (2005).
- Timmer, J. & Koenig, M. On generating power law noise. *Astron. Astrophys.* **300**, 707–710 (1995).
- Madejski, G. M. et al. Solving the mystery of the X-ray periodicity in the Seyfert galaxy NGC6814. *Nature* **365**, 626–628 (1993).
- Verbunt, F., Bunk, W. H., Ritter, H. & Pfeiffermann, E. Cataclysmic variables in the ROSAT PSPC all sky survey. *Astron. Astrophys.* **327**, 602–613 (1997).
- Bian, W. & Zhao, Y. Black hole masses in narrow-line Seyfert 1 galaxies. *Mon. Not. R. Astron. Soc.* **352**, 823–827 (2004).
- Kaspi, S. et al. Reverberation measurements for 17 quasars and the size–mass–luminosity relations in active galactic nuclei. *Astrophys. J.* **533**, 631–649 (2000).
- Tremaine, S. et al. The slope of the black hole mass versus velocity dispersion correlation. *Astrophys. J.* **574**, 740–753 (2002).
- Grupe, D. & Mathur, S. $M_{\text{BH}}-\sigma$ relation for a complete sample of soft X-ray-selected active galactic nuclei. *Astrophys. J.* **606**, L41–L44 (2004).
- Klein-Wolt, M. & van der Klis, M. Identification of black hole power spectral components across all canonical states. *Astrophys. J.* **675**, 1407–1423 (2008).
- Abramowicz, M. A. & Kluźniak, W. A precise determination of black hole spin in GRO J1655–40. *Astron. Astrophys.* **374**, L19–L20 (2001).
- Pounds, K. A., Done, C. & Osborne, J. P. RE 1034+39: A high-state Seyfert galaxy? *Mon. Not. R. Astron. Soc.* **277**, L5–L10 (1995).
- Casebeer, D. A., Leighly, K. M. & Baron, E. FUSE observation of the narrow-line Seyfert 1 galaxy RE 1034+39: Dependence of broad emission line strengths on the shape of the photoionizing spectrum. *Astrophys. J.* **637**, 157–182 (2006).
- Vaughan, S., Edelson, R., Warwick, R. S. & Uttley, P. On characterizing the variability properties of X-ray light curves from active galaxies. *Mon. Not. R. Astron. Soc.* **345**, 1271–1284 (2003).
- Papadakis, I. E. & Lawrence, A. Improved methods for power spectrum modelling of red noise. *Mon. Not. R. Astron. Soc.* **261**, 612–624 (1993).

Acknowledgements M.G. was supported in part by the Polish MNiSW (NN203065933). C.D. acknowledges financial support through a PPARC senior fellowship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.G. (marek.gierlinski@durham.ac.uk).

LETTERS

Unusual magnetic order in the pseudogap region of the superconductor $\text{HgBa}_2\text{CuO}_{4+\delta}$

Y. Li¹, V. Balédent², N. Barišić^{3,4}, Y. Cho^{3,5}, B. Fauqué², Y. Sidis², G. Yu¹, X. Zhao^{3,6}, P. Bourges² & M. Greven^{3,7}

The pseudogap region of the phase diagram is an important unsolved puzzle in the field of high-transition-temperature (high- T_c) superconductivity, characterized by anomalous physical properties^{1,2}. There are open questions about the number of distinct phases and the possible presence of a quantum-critical point underneath the superconducting dome^{3–5}. The picture has remained unclear because there has not been conclusive evidence for a new type of order. Neutron scattering measurements for $\text{YBa}_2\text{Cu}_3\text{O}_{6+\delta}$ (YBCO) resulted in contradictory claims of no^{6,7} and weak^{8,9} magnetic order, and the interpretation of muon spin relaxation measurements on YBCO^{10,11} and of circularly polarized photoemission experiments on $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$ (refs 12, 13) has been controversial. Here we use polarized neutron diffraction to demonstrate for the model superconductor $\text{HgBa}_2\text{CuO}_{4+\delta}$ (Hg1201) that the characteristic temperature T^* marks the onset of an unusual magnetic order. Together with recent results for YBCO^{14,15}, this observation constitutes a demonstration of the universal existence of such a state. The findings appear to rule out theories that regard T^* as a crossover temperature^{16–18} rather than a phase transition temperature^{19–21}. Instead, they are consistent with a variant of previously proposed charge-current-loop order^{19,20} that involves apical oxygen orbitals²², and with the notion that many of the unusual properties arise from the presence of a quantum-critical point^{3–5,19}.

YBCO has a relatively complicated orthorhombically distorted crystal structure, with two CuO_2 layers forming a double layer in the unit cell and with additional Cu–O chains between the double layers. Recent experiments^{14,15} on YBCO point to the possible existence of a new magnetic order below T^* (see Fig. 1a) that does not break translational invariance. The observed effect occurs at positions in reciprocal space that had not been considered in previous work. To assess whether this effect is unique to YBCO or a universal property of the high- T_c superconductors, it is essential to extend the investigation to additional, structurally simpler compounds. Hg1201 has a simple tetragonal crystal structure (Fig. 1e), with only one CuO_2 layer in the unit cell, and the highest maximum T_c of all known single-layer compounds^{23,24}. These properties, together with a wide accessible doping range and minimal effects of disorder^{24,25}, render Hg1201 an ideal system for the clarification of the pseudogap physics. Through recent advances in crystal growth, sizable high-quality single crystals have finally become available²⁶, making the present study possible.

Polarized neutron diffraction experiments were made on the 4F1 triple-axis spectrometer at the Laboratoire Léon Brillouin. The experimental setup was similar to that described previously¹⁴, allowing the detection of scattered neutrons in both spin-flip (SF) and non-spin-flip (NSF) channels. We define the flipping ratio

$\text{FR} = I_{\text{NSF}}/I_{\text{SF}}$ (where I is intensity) to quantify the experimental efficiency of polarization. With careful arrangement, a stable flipping ratio as high as 95 can be obtained. This proved to be crucial for the detection of small magnetic signals in samples with relatively high carrier concentration. All measurements were performed in the ($h0l$) scattering plane, where the scattering wave vector is quoted as $\mathbf{Q} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* \equiv (HKL)$ in units of the reciprocal lattice vectors, with typical room temperature values $a^* = b^* = 1.614 \text{ \AA}^{-1}$ and $c^* = 0.657 \text{ \AA}^{-1}$. Four underdoped samples were investigated; as revealed in Fig. 1b, they show sharp superconducting transitions with $T_c = 61, 79, 81$ and 89 K .

Figure 2a–c demonstrates the existence of a magnetic component in the spin-flip channel for samples A and B. Because of the relatively strong intensity from unavoidable nuclear Bragg peak leakage in the spin-flip geometry, the measurement was made at the weak nuclear reflection $\mathbf{Q} = (101)$. The neutron polarization was parallel to the momentum transfer, $\mathbf{P} \parallel \mathbf{Q}$, a geometry in which all magnetic scattering occurs in the spin-flip channel. The linear slope of the nuclear scattering observed in the non-spin-flip channel can be accounted for by the Debye–Waller factor. As expected, owing to the leakage, the spin-flip data have a linear nuclear scattering contribution as well. However, the spin-flip data also show an additional component below $T_{\text{mag}} \approx 250 \text{ K}$, which we conclude to be of magnetic origin (see also below). The two samples have nearly identical values of T_c and T_{mag} , and the strength of the magnetic signal is nearly indistinguishable after normalization by the nuclear scattering intensity (Fig. 2b,c). The onset of magnetic order in YBCO has been associated with the pseudogap temperature T^* , determined by resistivity measurements¹⁴. Resistivity data for a separate small crystal with $T_c = 79 \text{ K}$ are shown in Fig. 1c. Indeed, the rescaled magnetic intensities for samples A and B follow the deviation from linear resistivity quite well (Fig. 1d), strongly suggesting that the observed magnetic and charge properties share the same physical origin.

For new magnetic order associated with the pseudogap phase, it is expected that the ordering temperature and strength increase (decrease) towards lower (higher) doping. To test this, we subjected sample B to a reducing heat treatment that lowered the oxygen (and, consequently, the hole carrier) concentration. The resultant sample B' has a significantly lower T_c of 61 K . Indeed, as displayed in Fig. 2e,f, the onset of the magnetic order has shifted to significantly higher temperature, and the signal strength has increased by more than a factor of five. On the other hand, for the most highly doped crystal (sample C; $T_c = 89 \text{ K}$), we were no longer able to discern a magnetic signal within the counting statistics of the experiment (Fig. 2d).

Comparison between Hg1201 and YBCO^{14,15} demonstrates remarkable universality (Fig. 3). (i) In both cases, the order preserves the translational symmetry of the underlying lattice, unlike conventional

¹Department of Physics, Stanford University, Stanford, California 94305, USA. ²Laboratoire Léon Brillouin, CEA-CNRS, CEA-Saclay, 91191 Gif sur Yvette, France. ³Stanford Synchrotron Radiation Laboratory, Stanford, California 94309, USA. ⁴Physikalisches Institut, Universität Stuttgart, Pfaffenwaldring 57, 70550 Stuttgart, Germany. ⁵BK21 Team of Nano Fusion Technology, Pusan National University, Busan 609-735, Korea. ⁶State Key Lab of Inorganic Synthesis and Preparative Chemistry, College of Chemistry, Jilin University, 2699 Qianjin Street, Changchun 130012, China. ⁷Department of Applied Physics, Stanford University, Stanford, California 94305, USA.

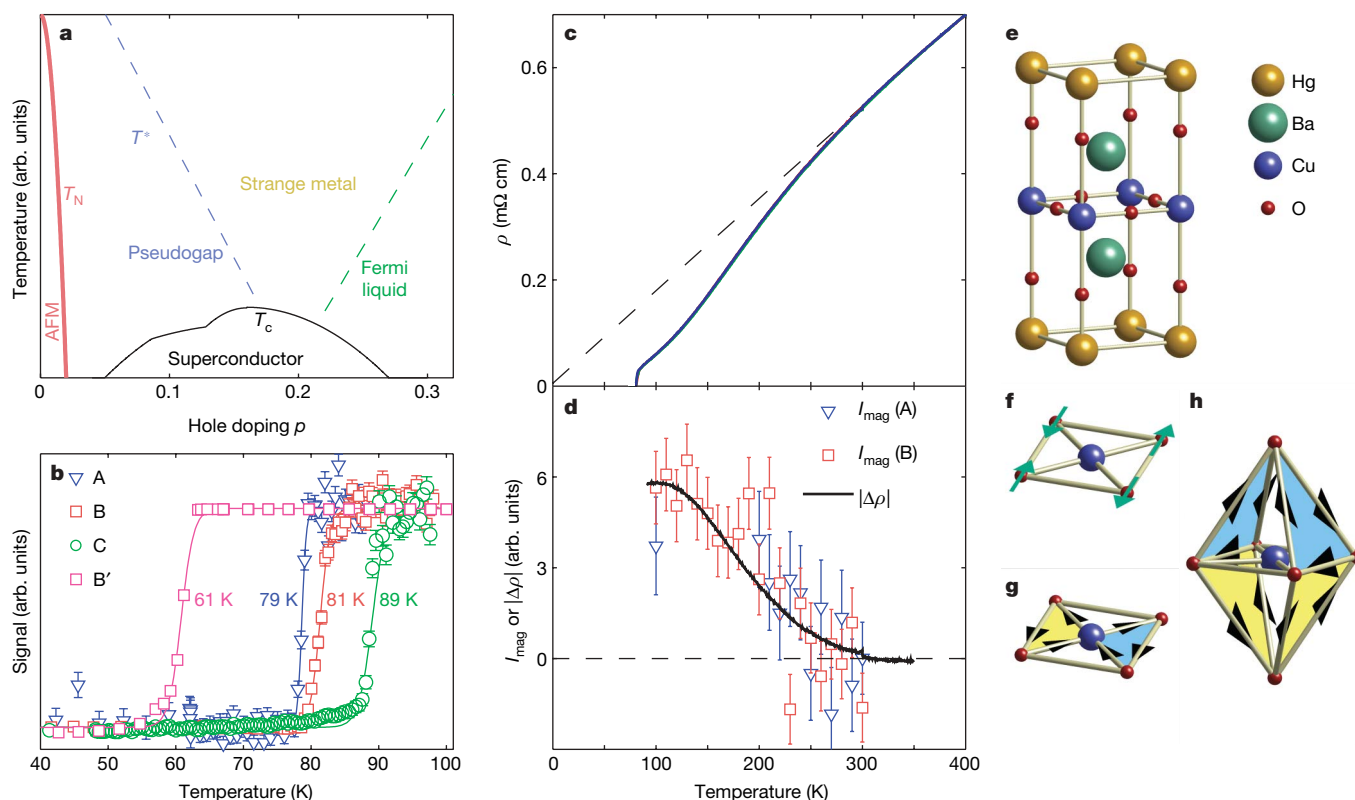


Figure 1 | Pseudogap in underdoped Hg1201. **a**, Schematic phase diagram of hole-doped high- T_c superconductors. T_N , Néel temperature of antiferromagnetism (AFM). **b**, Determination of T_c by superconducting quantum interference device (SQUID) magnetometry (for B') and neutron depolarization (for A, B, C) of the four underdoped Hg1201 samples studied in this work. Quoted T_c values are transition mid-points. Sample B' was obtained from sample B (mass ~ 600 mg, as-grown) by annealing in partial vacuum (0.1 torr) at 450°C . Samples A (150 mg) and C (1.2 g) are as-grown. Typical sample mosaic is less than 0.5° (full width at half maximum). The neutron-depolarization effect was used to measure the bulk T_c whenever possible. A guide field of ~ 10 Oe was applied along the beam path. After the samples were cooled below T_c and contained trapped vortices, the guide field at the sample position was turned by 90° , resulting in an abrupt change of magnetic field at the sample surface, which can be observed as a decrease in

flipping ratio. Error bars represent counting statistics (one standard deviation). In several cases, T_c was verified using conventional magnetic susceptibility measurements. **c**, Temperature dependence of resistivity for a separate crystal ($T_c = 79$ K). The resistivity measurement used the standard four-probe method, with electrical contacts sputtered on the ac/bc faces of a small single crystal (contact resistance less than 1 ohm). **d**, Deviation from linear resistivity compared with magnetic signal intensity (I_{mag}) for samples A and B, demonstrating that the new form of magnetic order is linked to the pseudogap. Error bars represent counting statistics (one standard deviation). **e**, Crystal structure of Hg1201 (dopant oxygen atoms in Hg–O layer not shown). **f–h**, Simplified schematic illustrations of three ordered states that break time-reversal symmetry, but preserve translational symmetry: **f**, spin-order involving oxygen atoms; **g**, planar orbital currents; **h**, orbital currents involving apical oxygens.

($\frac{1}{2}, \frac{1}{2}, 0$) type antiferromagnetism. (ii) The magnetic scattering develops below a temperature which coincides with T^* determined from d.c. transport, suggesting that the order involves both magnetic and charge degrees of freedom. (iii) The magnetic signal is of comparable strength for the two compounds, it is strongest in very underdoped samples, and the transition appears to be continuous. We note that the effect in the most underdoped sample B' is very strong, corresponding to $\sim 0.2\mu_B$ per unit cell in a naïve picture of spin-based moments, and that the present data statistics do not allow a reliable determination of the order parameter critical exponent. (iv) Using previous estimates^{27,28} for the doping dependence of T_c , the ordering temperatures for both systems fall onto the same line. Linear extrapolation suggests that T_{mag} approaches zero close to the value $p_c = 0.19$, which has been argued to be the location of a quantum critical point⁵. Alternatively, rescaling $T_c(p)$ for Hg1201 to the curve for YBCO shifts $T_{\text{mag}}(p)$ to higher hole concentrations and leads to an apparent disappearance of the magnetic signal near $p = 0.15$ for both systems (Supplementary Fig. 1). Both the linear trend and the value of p_c are consistent with new polar Kerr effect results for YBCO that also indicate the existence of a phase with broken time-reversal symmetry, although with ordering temperatures that are systematically lower²⁹. (v) In both cases, the moment does not lie along the c axis, but rather has a considerable in-plane component. For sample B, the (101) intensity measured with $\mathbf{P} \perp \mathbf{Q}$ (\mathbf{P} in the scattering plane) is about 65% of that for $\mathbf{P} \parallel \mathbf{Q}$ (Fig. 2c).

Noting that polarized neutron diffraction in the spin-flip channel probes the component of the magnetic moments perpendicular to both \mathbf{P} and \mathbf{Q} , and that in the former geometry \mathbf{P} makes a relatively small angle with the c axis, this suggests a non-negligible component of the measured moment in the a – b plane.

The observation of magnetic Bragg scattering at $\mathbf{Q} = (101)$ is consistent with an even number of moments per unit cell with zero net moment. Magnetic order involving spin moments on the planar oxygen atoms (Fig. 1f) could, in principle, preserve the translational invariance of the underlying crystal lattice¹⁴. However, such order would be difficult to reconcile with the unusual moment direction and, as discussed below, with the observed strong Q -dependence. Moreover, it should be discernable with NMR, yet no such evidence has been reported²⁵. Instead, it seems likely that the new state arises from circulating charge currents^{19,20}. The experiments for Hg1201 and YBCO^{14,15} are qualitatively consistent with magnetism due to two counter-circulating charge current loops per CuO_2 plaquette (Fig. 1g), but as the theory involves the planar oxygen p and copper d orbitals, it predicts a magnetic moment along the c axis, which cannot explain the in-plane component found experimentally. To explain the unusual moment direction and the tiny ferromagnetic component observed by the polar Kerr effect²⁹, it has been proposed that the relatively low structural symmetry of YBCO will lead to spin–orbit coupling that causes spin order to accompany planar loop-current

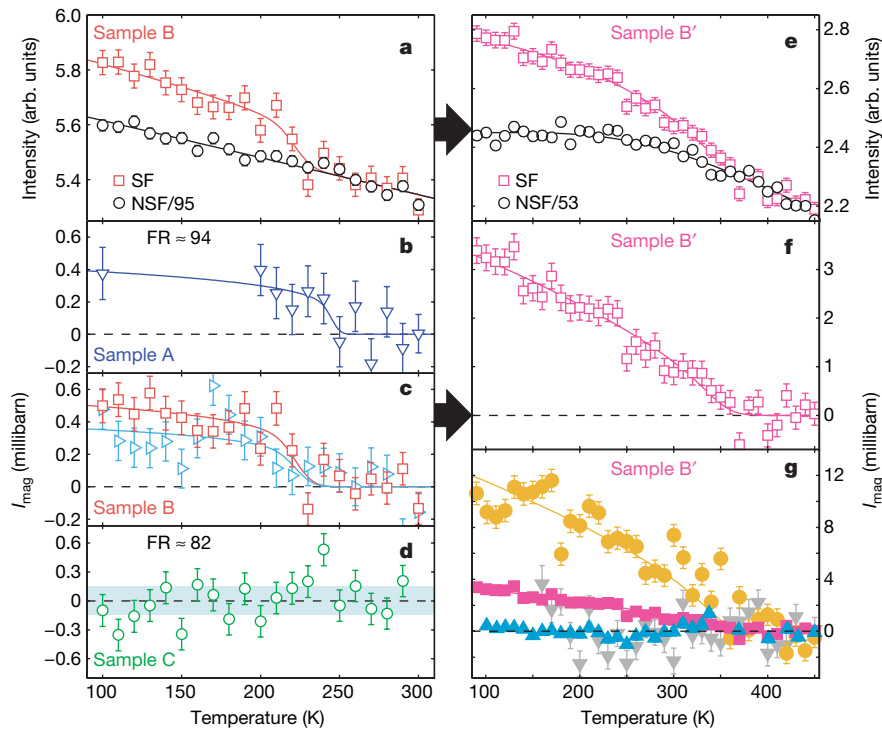


Figure 2 | Unusual magnetic order revealed by polarized-neutron diffraction. **a, e,** Raw data for samples B ($T_c = 81$ K) and B' ($T_c = 61$ K). Magnetic signal appears as additional intensity in the spin-flip (SF) channel compared with 'background' intensity due to nuclear Bragg scattering. The latter is measured in the non-spin-flip (NSF) channel, but a fraction $1/\text{FR}$ (where FR is the flipping ratio, such as 95 in 'NSF/95') is also observed in the spin-flip geometry. **b–d, f, g,** Temperature dependence of net intensity I_{mag} , which is obtained after the removal of the background. Sample A, $T_c = 79$ K; sample B, $T_c = 81$ K; sample C, $T_c = 89$ K. Conversion to absolute units is completed using the intensity of the same nuclear Bragg peak. **g,** Intensity measured on different Bragg peaks in the most underdoped sample. Yellow circles, (100); red squares, (101); blue triangles, (201); grey inverted triangles, (102). Data in **a–f** are collected on the Bragg peak $\mathbf{Q} = (101)$, with

the neutron spin parallel to \mathbf{Q} . In **c**, data are also collected with the neutron spin perpendicular to \mathbf{Q} in the scattering plane (red squares, $\mathbf{P} // \mathbf{Q}$; blue triangles, $\mathbf{P} \perp \mathbf{Q}$). Horizontal band in **d** represents an upper bound estimate for the intensity of sample C. Solid coloured lines in **a–c** and **e–g** are guides to the eye. Horizontal arrows represent the oxygen-anneal step carried out to obtain sample B'. Error bars represent counting statistics (one standard deviation). The polarized-neutron diffraction experiment was carried out in continuous runs with minimal instrument movement. The super-mirror polarizer and Heusler analyser were arranged such that the electromagnetic spin flipper was off when measuring the spin-flip channel, providing maximum stability. The experiments are limited to temperatures above T_c because the required neutron guide field cannot be reliably sustained in the superconducting state.

order³⁰. However, such spin–orbit coupling is expected to be absent in Hg1201, which possesses a high tetragonal structural symmetry in which the planar Cu and O sites are centres of inversion. The presence of significant oxygen spin moments is furthermore inconsistent with the narrow ^{17}O NMR linewidth²⁵.

The correct description might be a variant of the proposed phase, with orbital-current loops that involve the apical oxygens, but

without current flow through the copper site (Fig. 1h). The tilt angle spanned by the CuO_2 plane and the 'oxygen triangles' is about 64° for Hg1201 and 59° for YBCO, consistent with the fact that a large portion of the total signal ($\mathbf{P} // \mathbf{Q}$) is distributed in the $\mathbf{P} \perp \mathbf{Q}$ geometry

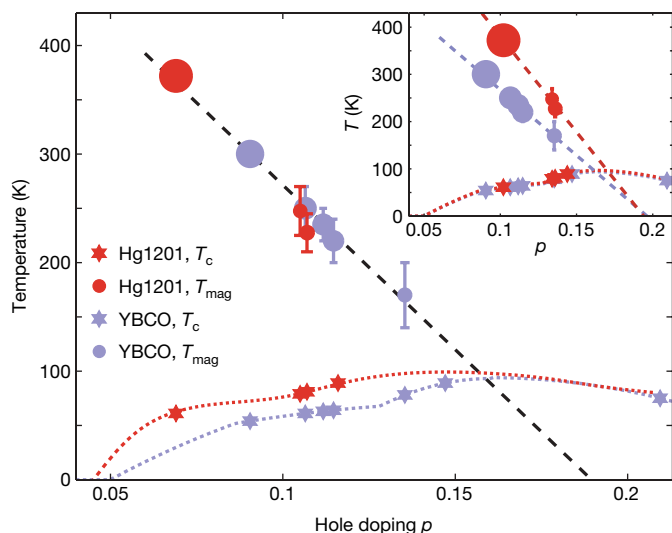


Figure 3 | Universal pseudogap phase diagram. In the main panel, hole doping is estimated from the $T_c(p)$ relationships (dotted lines) reported in^{27,28}. Values of T_{mag} for Hg1201 and YBCO are determined in this work and in refs 14,15, respectively. A linear fit of $T_{\text{mag}}(p)$ to the combined data extrapolates to $T_{\text{mag}} = 0$ K at $p_c = 0.190 \pm 0.011$ (black dashed line). Note that the value of T_c is a function not only of p but also of disorder^{24,27}, possibly leading to systematic differences in carrier concentration estimates. Furthermore, the $T_c(p)$ relationships for Hg1201 and YBCO differ below optimal doping, and the systematic deviation from a parabolic form might result from a tendency towards stripe-order formation near $p = 1/8$ (ref. 27). In the inset, p for Hg1201 is estimated using the $T_c(p)$ relationship for YBCO²⁷, and linear extrapolation gives $p_c = 0.194 \pm 0.025$ (Hg1201) and $p_c = 0.196 \pm 0.011$ (YBCO). Symbols are plotted with area proportional to signal intensity estimated at $T = 0$ K. Samples at relatively high doping did not show an observable magnetic signal (see also Supplementary Fig. 1 for YBCO). Error bars represent the uncertainty in the estimation of T_{mag} . The results are summarized in Table 1. For the most underdoped Hg1201 sample B', an order parameter fit was attempted, allowing for a small distribution of transition temperatures T_{mag} . This distribution was estimated from the superconducting transition width and the approximate change of T_{mag} with T_c . Fits to the (100) and (101) Bragg peaks were carried out to extract T_{mag} and the exponent β ; β was found to be strongly dependent on the range of the fit, which was unstable for $T > 300$ K, and yielded $\beta = 0.18 \pm 0.13$ for $T > 250$ K and $\beta = 0.42 \pm 0.12$ for $T > 200$ K.

Table 1 | Summary of results for the universal pseudogap phase diagram

<i>P</i>	<i>T_c</i> (K)	<i>T_{mag}</i> (K)	<i>I_{mag}</i> (mb)
Y 0.091	54	300 ± 10	2.8 ± 0.3
Hg 0.069 (0.102)	61	372 ± 13	4.2 ± 0.3
Y 0.107	61	250 ± 20	1.7 ± 0.2
Y 0.112	63	235 ± 15	1.6 ± 0.2
Y 0.115	64	220 ± 20	1.5 ± 0.2
Y 0.135	78	170 ± 30	0.6 ± 0.1
Hg 0.105 (0.134)	79	248 ± 23	0.5 ± 0.1
Hg 0.107 (0.136)	81	228 ± 18	0.6 ± 0.1
Hg 0.116 (0.144)	89	N/A	<0.15
Y 0.147	89	N/A	<0.2
Y 0.209	75	N/A	<0.1

The hole concentration *p* estimated from the *T_c*(*p*) relationships reported previously^{27,28} (values for Hg1201 in parentheses are those pertaining to the inset to Fig. 3), superconducting transition temperature, onset temperature and strength of magnetic signal reported in this work and in^{14,15}. Hg and Y denote Hg1201 and YBCO, respectively. The data are shown in Fig. 3. Neutron diffraction data for the YBCO 0.147 sample can be found in Supplementary Fig. 1.

(**P** in the scattering plane). These observations are further supported by the data in Fig. 2g, which demonstrate that the magnetic signal is even stronger at **Q** = (100) than at (101), whereas at the (201) and (102) reflections no signal was discerned. This trend, which is in good agreement with the results for YBCO¹⁴, is consistent with the general expectation that magnetic signal decreases with increasing **Q**. However, the sharp decrease of the intensity with increasing **Q** implies that, distinct from conventional spin order, the magnetic density has a large spatial extent, consistent with a picture of extended spontaneous orbital currents within the unit cell. Moreover, the currents cannot be confined to the CuO₂ planes (Fig. 1g), as this would not lead to a strong *L*-dependence. Very recent theoretical work on extended two-dimensional Hubbard models including apical oxygen orbitals supports this picture²².

The maximum *T_c* occurs close to where the experiment fails to discern a magnetic signal, and it seems likely that the order competes with the superconductivity. One intriguing possibility is that the fluctuations associated with an underlying quantum critical point are directly responsible for the appearance of superconductivity and the unusual normal state properties, such as the linear resistivity found up to remarkably high temperatures.

Received 14 May; accepted 2 July 2008.

- Timusk, T. & Statt, B. The pseudogap in high-temperature superconductors: an experimental survey. *Rep. Prog. Phys.* **62**, 61–122 (1999).
- Norman, M. R., Pines, D. & Kallin, C. The pseudogap: friend or foe of high *T_c*? *Adv. Phys.* **54**, 715–733 (2005).
- Laughlin, R. B. A critique of two metals. *Adv. Phys.* **47**, 943–958 (1998).
- Sachdev, S. Quantum criticality: competing ground states in low dimensions. *Science* **288**, 475–480 (2000).
- Tallon, J. L. & Loram, J. W. The doping dependence of *T^{*}*: what is the real high-*T_c* phase diagram? *Physica C* **349**, 53–68 (2001).
- Lee, S.-H. et al. Search for orbital moments in underdoped cuprate metals. *Phys. Rev. B* **60**, 10405–10417 (1999).

- Stock, C. et al. Neutron scattering search for static magnetism in oxygen-ordered YBa₂Cu₃O_{6.5}. *Phys. Rev. B* **66**, 024505 (2002).
- Sidis, Y. et al. Antiferromagnetic ordering in superconducting YBa₂Cu₃O_{6.5}. *Phys. Rev. Lett.* **86**, 4100–4103 (2001).
- Mook, H. A. et al. Polarized neutron measurement of magnetic order in YBa₂Cu₃O_{6.45}. *Phys. Rev. B* **69**, 134509 (2004).
- Sonier, J. E. et al. Anomalous weak magnetism in superconducting YBa₂Cu₃O_{6+x}. *Science* **292**, 1692–1695 (2001).
- Sonier, J. E. et al. Correlations between charge ordering and local magnetic fields in overdoped YBa₂Cu₃O_{6+x}. *Phys. Rev. B* **66**, 134501 (2002).
- Kaminski, A. et al. Spontaneous breaking of time-reversal symmetry in the pseudogap state of a high-*T_c* superconductor. *Nature* **416**, 610–613 (2002).
- Borisenko, S. V. et al. Circular dichroism in angle-resolved photoemission spectra of under- and overdoped Pb-Bi2212. *Phys. Rev. Lett.* **92**, 207001 (2004).
- Fauqué, B. et al. Magnetic order in the pseudogap phase of high-*T_c* superconductors. *Phys. Rev. Lett.* **96**, 197001 (2006).
- Mook, H. A. et al. Observation of magnetic order in a YBa₂Cu₃O_{6.6} superconductor. *Phys. Rev. B* **78**, 020506 (2008).
- Emery, V. J. & Kivelson, S. A. Importance of phase fluctuations in superconductors with small superfluid density. *Nature* **374**, 434–437 (1995).
- Lee, P. A. Pseudogaps in underdoped cuprates. *Physica C* **317**, 194–204 (1999).
- Anderson, P. W. et al. The physics behind high-temperature superconducting cuprates: the 'plain vanilla' version of RVB. *J. Phys. Cond. Mater.* **16**, R755–R769 (2004).
- Varma, C. M. Non-Fermi-liquid states and pairing instability of a general model of copper oxide metals. *Phys. Rev. B* **55**, 14554–14580 (1997).
- Varma, C. M. Theory of the pseudogap state of the cuprates. *Phys. Rev. B* **73**, 155113 (2006).
- Chakravarty, S., Laughlin, R. B., Morr, D. K. & Nayak, C. Hidden order in the cuprates. *Phys. Rev. B* **63**, 094503 (2001).
- Weber, C., Läuchli, A., Mila, F. & Giamarchi, T. Orbital currents in extended Hubbard models of high-*T_c* cuprates. Preprint available at (<http://arxiv.org/abs/0803.3983>).
- Putlin, S. N., Antipov, E. V., Chmaissem, O. & Marezio, M. Superconductivity at 94 K in HgBa₂CuO_{4+δ}. *Nature* **362**, 226–228 (1993).
- Eisaki, H. et al. Effect of chemical inhomogeneity in bismuth-based copper oxide superconductors. *Phys. Rev. B* **69**, 064512 (2004).
- Bobroff, J. et al. ¹⁷O NMR evidence for a pseudogap in the monolayer HgBa₂CuO_{4+δ}. *Phys. Rev. Lett.* **78**, 3757–3760 (1997).
- Zhao, X. et al. Crystal growth and characterization of the model high-temperature superconductor HgBa₂CuO_{4+δ}. *Adv. Mater.* **18**, 3243–3247 (2006).
- Liang, R., Bonn, D. A. & Hardy, W. N. Evaluation of CuO₂ plane hole doping in YBa₂Cu₃O_{6+x} single crystals. *Phys. Rev. B* **73**, 180505 (2006).
- Yamamoto, A. et al. Thermoelectric power and resistivity of HgBa₂CuO_{4+δ} over a wide doping range. *Phys. Rev. B* **63**, 024504 (2001).
- Xia, J. et al. Polar Kerr-effect measurement of the high-temperature YBa₂Cu₃O_{6+x} superconductor: evidence for broken symmetry near the pseudogap temperature. *Phys. Rev. Lett.* **100**, 127002 (2008).
- Aji, V. & Varma, C. M. Spin order accompanying loop-current order in cuprate superconductors. *Phys. Rev. B* **75**, 224511 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank H. Alloul and C. Varma for comments. The work at Stanford University was supported by grants from the US Department of Energy and the National Science Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.G. (greven@stanford.edu).

LETTERS

Three-dimensional optical metamaterial with a negative refractive index

Jason Valentine^{1*}, Shuang Zhang^{1*}, Thomas Zentgraf^{1*}, Erick Ulin-Avila¹, Dentcho A. Genov¹, Guy Bartal¹ & Xiang Zhang^{1,2}

Metamaterials are artificially engineered structures that have properties, such as a negative refractive index^{1–4}, not attainable with naturally occurring materials. Negative-index metamaterials (NIMs) were first demonstrated for microwave frequencies^{5,6}, but it has been challenging to design NIMs for optical frequencies and they have so far been limited to optically thin samples because of significant fabrication challenges and strong energy dissipation in metals^{7,8}. Such thin structures are analogous to a monolayer of atoms, making it difficult to assign bulk properties such as the index of refraction. Negative refraction of surface plasmons was recently demonstrated but was confined to a two-dimensional waveguide⁹. Three-dimensional (3D) optical metamaterials have come into focus recently, including the realization of negative refraction by using layered semiconductor metamaterials and a 3D magnetic metamaterial in the infrared frequencies; however, neither of these had a negative index of refraction^{10,11}. Here we report a 3D optical metamaterial having negative refractive index with a very high figure of merit of 3.5 (that is, low loss). This metamaterial is made of cascaded ‘fishnet’ structures, with a negative index existing over a broad spectral range. Moreover, it can readily be probed from free space, making it functional for optical devices. We construct a prism made of this optical NIM to demonstrate negative refractive index at optical frequencies, resulting unambiguously from the negative phase evolution of the wave propagating inside the metamaterial. Bulk optical metamaterials open up prospects for studies of 3D optical effects and applications associated with NIMs and zero-index materials such as reversed Doppler effect, superlenses, optical tunnelling devices^{12,13}, compact resonators and highly directional sources¹⁴.

NIMs, first described by Veselago more than 40 years ago¹ and recently discussed in the framework of metamaterials², arise from the fact that both the permittivity and the permeability of the materials are simultaneously negative. In the past several years, much effort has been dedicated to the engineering and extension of the functionalities of metamaterials at terahertz^{15–17} and optical frequencies^{7,8,10,18–21}. Metal–dielectric–metal fishnet structures were among the earliest demonstrations of optical NIMs. These structures, however, consist of a single functional layer along the direction of propagation. This is equivalent to an atomic monolayer, making it difficult to explore phenomena in three dimensions and develop device applications. Moreover, as a result of their resonant nature, these systems suffer substantial loss at optical frequencies. On the basis of the above, it is therefore imperative to realize low-loss bulk optical NIMs if we are to demonstrate unambiguously the unique effects associated with negative index of refraction.

Recently, it has been suggested theoretically that stacking up multiple fishnet functional layers along the propagation direction

constitutes a promising approach for achieving a 3D optical NIM²² (Fig. 1a). This cascading leads to a strong magneto-inductive coupling between neighbouring functional layers²³. As demonstrated recently²⁴, the tight coupling between adjacent LC resonators through mutual inductance results in a broadband negative index of refraction with low loss, which is similar to the material response of left-handed transmission lines^{25,26}. In addition, the loss is further reduced owing to the destructive interference of the antisymmetric currents across the metal film, effectively cancelling out the current flow in the centre of the film²³.

Here we experimentally demonstrate the first 3D optical NIM by directly measuring the angle of refraction from a prism made of cascaded fishnet metamaterial. The experimental results, along with numerical calculations, serve as direct evidence of zero and negative phase index in the metamaterial.

The 3D fishnet metamaterial is fabricated on a multilayer metal–dielectric stack by using focused ion-beam milling (FIB), which is capable of cutting nanometre-sized features with a high aspect ratio. Figure 1b shows the scanning electron microscopy (SEM) image of the proposed 3D fishnet pattern, which was milled on 21 alternating films of silver and magnesium fluoride, resulting in ten functional layers.

To measure the index of refraction of the 3D metamaterial experimentally, a prism was created in the multilayer stack (Fig. 2a, b). Measurements of the refractive index of these structures were performed by observing the refraction angle of light passing through the prism by Snell’s law. This provides a direct and unambiguous

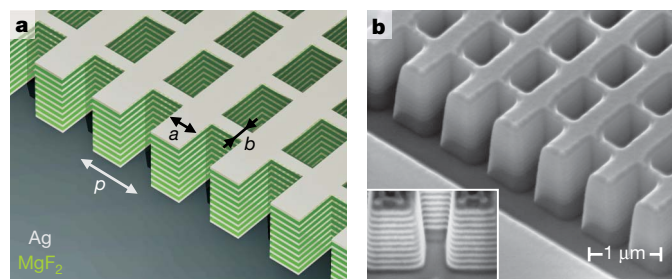


Figure 1 | Diagram and SEM image of fabricated fishnet structure. **a**, Diagram of the 21-layer fishnet structure with a unit cell of $p = 860$ nm, $a = 565$ nm and $b = 265$ nm. **b**, SEM image of the 21-layer fishnet structure with the side etched, showing the cross-section. The structure consists of alternating layers of 30 nm silver (Ag) and 50 nm magnesium fluoride (MgF_2), and the dimensions of the structure correspond to the diagram in **a**. The inset shows a cross-section of the pattern taken at a 45° angle. The sidewall angle is 4.3° and was found to have a minor effect on the transmittance curve according to simulation.

¹NSF Nano-scale Science and Engineering Center (NSEC), 3112 Etcheverry Hall, University of California, Berkeley, California 94720, USA. ²Material Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA.

*These authors contributed equally to this work.

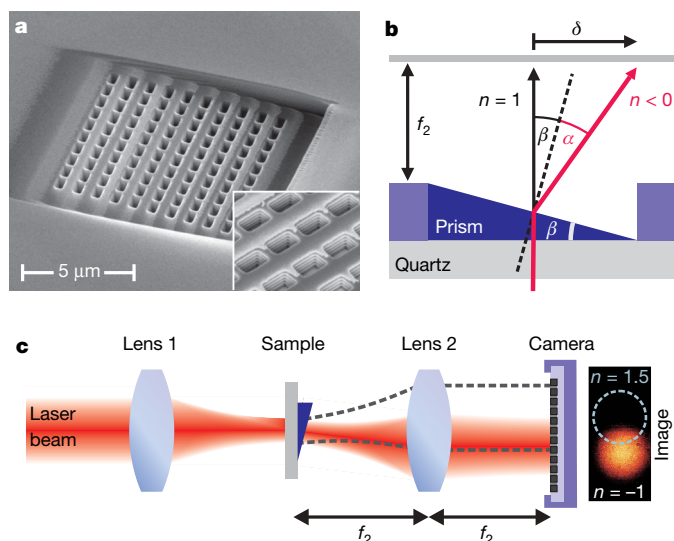


Figure 2 | SEM image of NIM prism and schematics of experimental setup. **a**, SEM image of the fabricated 3D fishnet NIM prism. The unit cell size is identical to that shown in Fig. 1a. The inset shows a magnified view with the film layers visible in each hole. **b**, Geometry diagram of the angle measurement; δ corresponds to the position difference of the beam passing through a window in the multilayer structure ($n = 1$) and prism sample. By measuring δ , the absolute angle of refraction α can be obtained. **c**, Experimental setup for the beam refraction measurement. The focal length of lens 1 is 50 mm and that of lens 2 is $f_2 = 40$ mm. Lens 2 is placed in a $2f$ configuration, resulting in the Fourier image at the camera position.

determination of the refractive index, because the refraction angle depends solely on the phase gradient that the light beam experiences when refracted from the angled output face. We used a femtosecond synchronously pumped optical parametric oscillator as a tunable

light source to determine the refractive index at different wavelengths. The beam was focused on the sample, and a charge-coupled device (CCD) camera was placed in the Fourier plane (Fig. 2c).

Figure 3a shows the beam shift δ resulting from the bending of light at the prism output at different wavelengths, ranging from 1,200 to 1,700 nm. The measurement was performed on a prism of angle $\beta = 5.0^\circ$ and the beam shift is plotted along with reference measurements of transmission through a window, without the presence of the prism (left panel). Clearly, as the wavelength increases, the beam shift resulting from the prism refraction is changing from positive to negative, indicating a transition from a positive index in the shorter wavelengths to a negative index in the longer wavelengths. At a wavelength λ of 1,475 nm, the index of refraction is approaching zero; that is, the beam does not acquire any phase while propagating in the metamaterial. Consequently there is no phase gradient at the angled output face and the exiting beam is exactly normal to the output face (see dashed lines in Figs 2b and 3a).

Figure 3b depicts the measured refractive index of the 3D fishnet metamaterial at various wavelengths. The refractive index varies from $n = 0.63 \pm 0.05$ at 1,200 nm to $n = -1.23 \pm 0.34$ at 1,775 nm. The refractive index was determined from multiple measurements of two fishnet prisms with angles of $\beta = 5.0^\circ$ and 4.7° and for wavelengths ranging from 1,200 to 1,800 nm. Although there is a correlation between the beam spot positions shown in Fig. 3a and the refractive index in Fig. 3b, it should be noted that Fig. 3b shows the average of measurements on different prisms with the standard deviation as error bars, whereas Fig. 3a shows an individual measurement. The experimental results are found to be in good agreement with the theoretical predictions (black line in Fig. 3b) on the basis of rigorous coupled-wave analysis (RCWA). The measured negative refraction angle is a direct result of negative phase evolution for light propagating inside the sample caused by a negative refractive index. This is illustrated in Fig. 3c by a numerical calculation of the in-plane electromagnetic field distribution in the fishnet prism at

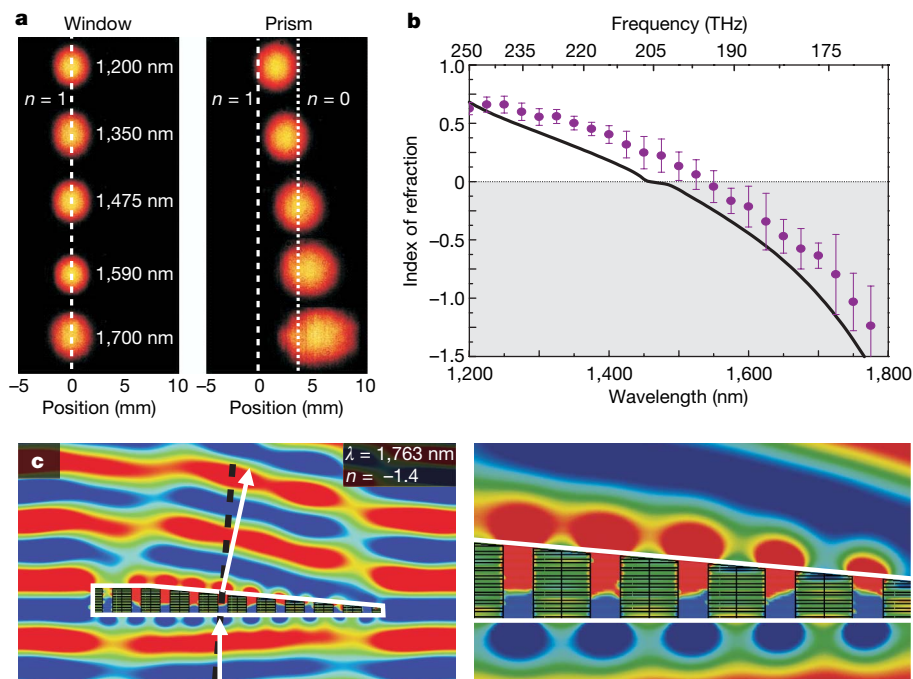


Figure 3 | Experimental results and finite-difference time-domain simulations. **a**, Fourier-plane images of the beam for the window and prism sample for various wavelengths. The horizontal axis corresponds to the beam shift δ , and positions of $n = 1$ and $n = 0$ are denoted by the white lines. The image intensity for each wavelength has been normalized for clarity. **b**, Measurements and simulation of the fishnet refractive index. The circles show the results of the experimental measurement with error bars (s.d.,

$n = 4$ measurements). The measurement agrees closely with the simulated refractive index using the RCWA method (black line). **c**, Left: simulation of the in-plane electric field component for the prism structure at 1,763 nm, showing the phase front of the light. Negative-phase propagation resulting from the negative refractive index leads to negative refraction angle as measured by the beam shift in the experiment. Right: magnified plot of the field distribution in the prism.

$\lambda = 1,763$ nm, where the structure shows a refractive index of $n = -1.4$. A movie of the evolution of the electric field generated with commercial finite-difference time-domain software at the same wavelength is also presented in Supplementary Information. Because of the negative phase propagation inside the metamaterial, the electromagnetic wave emerging from the thicker part of the prism experiences phase advance compared with that passing through the thinner parts, causing the light to bend in the negative direction at the exit interface. We note that the refractive index remains consistent for the fishnet metamaterial with three or more functional layers along the propagation direction, which leads to a uniform wavefront exiting from the prism (see Supplementary Information).

To acquire a clear understanding of the 3D metamaterial's optical response, we separate the fishnet into two subsets and calculate the dispersion curves with RCWA. The first constituent is a 3D array of metal wires aligned with the polarization direction of the incident electric field (Fig. 4a). This array serves as an effective medium with lowered volumetric plasma frequency (220 THz), below which wave propagation is not allowed because of negative effective permittivity²⁷. The second constituent is a 3D array of metal strips along the direction of the magnetic field (Fig. 4b), in which induced antisymmetric conductive currents across the dielectric layers give rise to a magnetic bandgap between 135 and 210 THz. This is further confirmed by the plots of the magnetic fields at two frequencies, below and above the bandgap. Above the bandgap the magnetic response is positive, as shown in Fig. 4d, where the magnetic field component between the strips is in phase with the external field. Above the

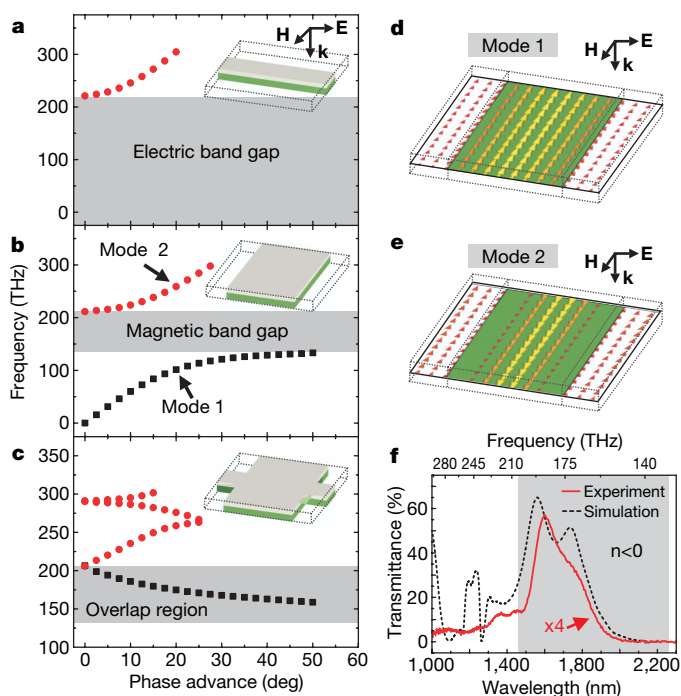


Figure 4 | Dispersion relations, field plots and transmission for the 3D fishnet structure. In all plots the grey area corresponds to the negative-index region as determined by simulations. **a**, Dispersion relation for a 3D array of metal wires aligned along the electric field \mathbf{E} , where \mathbf{k} denotes the incident propagation vector. The dotted lines in the diagram mark the unit cell size, which is identical in **a–e**, and correspond to Fig. 1a. **b**, Dispersion relation of a 3D array of metal strips along the magnetic field \mathbf{H} . **c**, The dispersion for the 3D fishnet structure. A dispersion curve with negative slope appears within the overlapped region of the electric bandgap and magnetic bandgap if both structures are combined. **d**, Magnetic field plot below the bandgap. **e**, Magnetic field plot above the bandgap. **f**, Experimental transmittance curve (red line) of a 22×22 unit cell fishnet structure ($17.6 \times 17.6 \mu\text{m}^2$ total patterned area), which has been multiplied by four for clarity. The simulated transmittance (black dotted line) was obtained with RCWA.

bandgap, the metal strips have a moderately negative response, as shown in Fig. 4e, where the magnetic field between the strips is out of phase. Finally, these two structures are merged to form the 3D fishnet metamaterial, for which the dispersion relation is shown in Fig. 4c. A propagation band with negative slope appears in the overlapped region of the forbidden gaps of both electric and magnetic media, demonstrating that the negative-index behaviour in the 3D cascaded fishnet does indeed result from the fact that both the electric permittivity and the magnetic permeability are negative.

In addition, transmittance measurements were performed on the 3D fishnet metamaterial made of 21 layers with the use of a Fourier-transform infrared microscope (Nicolet Nic-Plan IR microscope). Figure 4f shows the measured transmittance spectrum along with the numerically calculated transmittance. The simulation predicts a broad negative-index band spanning from $1.45 \mu\text{m}$ to $2.2 \mu\text{m}$ (shaded region), which coincides with the high transmission band from $1.5 \mu\text{m}$ to $1.8 \mu\text{m}$. As mentioned previously, this wide band of negative index results from the strong coupling between neighbouring layers. The measured transmittance has similar features to those of the calculation, namely two peaks imposed over the transmission band that are slightly red-shifted with respect to the numerical results. These features are due to the Fabry–Pérot effect, in which the impedance mismatch leads to reflectance at the metamaterial/air and metamaterial/glass interfaces. Although the peaks are visible at lower refractive index values owing to the lower loss, the Fabry–Pérot effect cannot be clearly seen in the transmission spectra for larger negative index at longer wavelength where the loss is higher, resulting in broadening and extinction of spectral features. We note that the transmittance in the negative-index band is one-quarter of the numerically calculated value, which is probably due to imperfections in the fabrication. Nevertheless, our simulations show that the presence of loss in the coupled fishnet metamaterial has a minimal impact on the dispersion relation (see Supplementary Information). This is because the 3D fishnet structure operates far from the band edge (Fig. 4c), where resonance is not significant. This explains the good agreement between the experimentally measured and simulated refractive indices despite the fabrication imperfections.

Finally, we estimate the figure of merit ($\text{FOM} = -\text{Re}(n)/|\text{Im}(n)|$) of our fishnet multilayer structure. The material loss (that is, $\text{Im}(n)$) is conservatively estimated from the measured transmittance and reflectance data of the 21-layer sample, assuming a single pass of light through the metamaterial, as $\text{Im}(n) = (\lambda/4\pi d) \ln[(1 - R)/T]$, where λ , d , R and T are the wavelength, sample thickness, reflectance and transmittance, respectively. The dispersion of the simulated

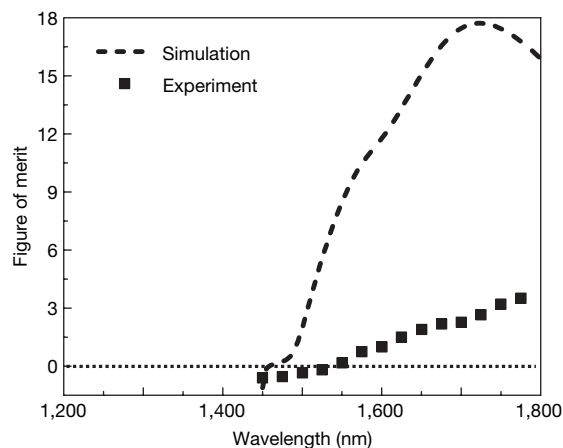


Figure 5 | Figure of merit for the 3D fishnet structure. Plot of FOM against wavelength for the simulation (dashed line) and experiment (squares). The lower experimental FOM is due to reduced transmission resulting from fabrication imperfections. The experimental FOM reaches 3.5 at 1,775 nm, where $\text{Re}(n) = -1.23$.

and experimental FOM is plotted in Fig. 5. The FOM is 3.5 at $\lambda = 1,775$ nm (where $\text{Re}(n) = -1.23$), which is among the highest experimental values so far recorded at optical frequencies¹⁹. For ideal fabrication conditions, the FOM could rise as high as about 20, as determined from theoretical calculations. We emphasize that our results are different from recent reports of negative refraction^{11,28} in anisotropic media with hyperbolic dispersion (equivalent to negative 'group index') but positive phase velocity.

The fishnet metamaterial has a period about $\lambda/20$ in the vertical direction. The propagation of light travelling along this direction or within some angular range is dominated by this deep sub-wavelength period and not by the in-plane period, as long as the wavevector projection on the in-plane directions is small compared with the in-plane reciprocal lattice vector of the fishnet metamaterial. There is only a single propagating mode in the negative-index frequency region, justifying the description of the fishnet metamaterial with an effective index. In contrast, if higher dielectric materials such as silicon ($n \approx 3.6$) are used to serve as the dielectric layer, the ratio between the wavelength and in-plane period can be significantly increased because of the larger capacitance in the LC circuit.

Unlike the negative index obtained from photonic crystals²⁹, the negative index presented here results from simultaneous negative magnetic and electric responses and shows a resemblance to the left-handed transmission line due to the tight coupling between the adjacent LC resonators. The negative index occurs in the first propagation band and with smooth negative-phase evolution along the light propagation direction, which differs from the negative refraction obtained in photonic crystals.

Here we have experimentally demonstrated the first 3D NIM at optical frequencies and directly measured the refractive index of a NIM prism in the free space. The 3D optical metamaterials may offer the possibility to explore a large variety of optical phenomena associated with zero and negative refractive index, as well as applications in the scaling down of photonics and imaging.

METHODS SUMMARY

In the numerical studies of the 3D fishnet metamaterial, the intrinsic losses of the metal are included³⁰. The multilayer stack was deposited by electron-beam evaporation of alternating layers of silver (30 nm) and magnesium fluoride (50 nm) resulting in a total thickness of 830 nm. Two different configurations of the fishnet samples were fabricated on the multilayer stack. Samples of the first configuration consist of 22×22 in-plane fishnet unit cells and were used for the characterization of the transmittance. The second configuration (prism sample) was formed by etching the film at an angle β to the film surface, using FIB. The exact angle was measured with an atomic force microscope and was found to vary slightly between samples. A 10×10 fishnet pattern was subsequently milled in the prism.

To obtain the absolute angle of refraction, a window with an area equal to that of the prism was etched through the multilayer stack to serve as a reference. The window and prism Fourier images were measured for all wavelengths on an indium gallium arsenide infrared camera and the total beam shift δ of the spot centre was calculated. Consequently, the angle of refraction α at the surface of the prism is given as $\alpha = \beta - \arctan(\delta/f_2)$. Snell's law ($n = \sin\alpha/\sin\beta$) was used to calculate the real part of the refractive index of the sample. The imaginary part of the refractive index of the sample was obtained from transmittance and reflectance data acquired with a 21-layer sample of the first configuration (as described above).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 March; accepted 11 July 2008.
Published online 11 August 2008.

1. Veselago, V. G. The electrodynamics of substances with simultaneously negative values of ϵ and μ . *Sov. Phys. Usp.* **10**, 509–514 (1968).
2. Smith, D. R., Pendry, J. B. & Wiltshire, M. C. K. Metamaterials and negative refractive index. *Science* **305**, 788–792 (2004).
3. Pendry, J. B. Negative refraction makes a perfect lens. *Phys. Rev. Lett.* **85**, 3966–3969 (2000).

4. Tsakmakidis, K. L., Boardman, A. D. & Hess, O. 'Trapped rainbow' storage of light in metamaterials. *Nature* **450**, 397–401 (2007).
5. Shelby, R. A., Smith, D. R. & Schultz, S. Experimental verification of a negative index of refraction. *Science* **292**, 77–79 (2001).
6. Parazzoli, C. G., Greengard, K., Li, K., Koltentbah, B. E. C. & Tanielian, M. Experimental verification of negative index of refraction using Snell's law. *Phys. Rev. Lett.* **90**, 107401 (2003).
7. Panoiu, N. C. & Osgood, R. M. Numerical investigations of negative refractive index metamaterials at infrared and optical frequencies. *Opt. Commun.* **223**, 331–337 (2003).
8. Shalaev, V. M. et al. Optical negative-index metamaterials. *Nature Photonics* **1**, 41–48 (2007).
9. Lezec, H. J., Dionne, N. A. & Atwater, H. A. Negative refraction at visible frequencies. *Science* **316**, 430–432 (2007).
10. Liu, N. et al. Three-dimensional photonic metamaterials at optical frequencies. *Nature Mater.* **7**, 31–37 (2008).
11. Hoffman, A. J. et al. Negative refraction in semiconductor metamaterials. *Nature Mater.* **6**, 946–950 (2007).
12. Silveirinha, M. & Engheta, N. Tunneling of electromagnetic energy through subwavelength channels and bends using epsilon-near-zero materials. *Phys. Rev. Lett.* **97**, 157403 (2006).
13. Edwards, B. et al. Experimental verification of epsilon-near-zero metamaterial coupling and energy squeezing using a microwave waveguide. *Phys. Rev. Lett.* **100**, 033903 (2008).
14. Enoch, S. et al. A metamaterial for directive emission. *Phys. Rev. Lett.* **89**, 213902 (2002).
15. Yen, T. J. et al. Terahertz magnetic response from artificial materials. *Science* **303**, 1494–1496 (2004).
16. Padilla, W. J. et al. Dynamical electric and magnetic metamaterial response at terahertz frequencies. *Phys. Rev. Lett.* **96**, 107401 (2006).
17. Chen, H. T. et al. Active terahertz metamaterial devices. *Nature* **444**, 597–600 (2006).
18. Linden, S. et al. Magnetic response of metamaterials at 100 terahertz. *Science* **306**, 1351–1353 (2004).
19. Soukoulis, C. M., Linden, S. & Wegener, M. Negative refractive index at optical frequencies. *Science* **315**, 47–49 (2007).
20. Dolling, G., Wegener, M. & Linden, S. Realization of a three-functional-layer negative-index photonic metamaterial. *Opt. Lett.* **32**, 551–553 (2007).
21. Alu, A. & Engheta, N. Three-dimensional nanotransmission lines at optical frequencies: A recipe for broad band negative-refraction optical metamaterials. *Phys. Rev. B* **75**, 024304 (2007).
22. Zhang, S. et al. Optical negative-index bulk metamaterials consisting of 2D perforated metal-dielectric stacks. *Opt. Express* **14**, 6778–6787 (2006).
23. Li, T. et al. Coupling effect of magnetic polariton in perforated metal/dielectric layered metamaterials and its influence on negative refraction transmission. *Opt. Express* **14**, 11155–11163 (2006).
24. Eleftheriades, G. V. Analysis of bandwidth and loss in negative-refractive-index transmission-line (NRI-TL) media using coupled resonators. *IEEE Microw. Wireless Components Lett.* **17**, 412–414 (2007).
25. Grbic, A. & Eleftheriades, G. V. Overcoming the diffraction limit with a planar left-handed transmission-line lens. *Phys. Rev. Lett.* **92**, 117403 (2004).
26. Lai, A., Carloz, C. & Itoh, T. Composite right-/left-handed composite transmission line metamaterials. *IEEE Microw. Mag.* **5**, 34–50 (2004).
27. Pendry, J. B., Holden, A. J., Robbins, D. J. & Stewart, W. J. Low frequency plasmons in thin-wire structures. *J. Phys. Condens. Matter* **10**, 4785–4809 (1998).
28. Fan, X. B. et al. All-angle broadband negative refraction of metal waveguide arrays in the visible range: Theoretical analysis and numerical demonstration. *Phys. Rev. Lett.* **97**, 073901 (2006).
29. Notomi, M. Theory of light propagation in strongly modulated photonic crystals: Refraction-like behavior in the vicinity of the photonic band gap. *Phys. Rev. B* **62**, 10696 (2000).
30. Johnson, P. B. & Christy, R. W. Optical constants of the noble metals. *Phys. Rev. B* **6**, 4370–4379 (1972).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge funding support from US Army Research Office (ARO) MURI programme 50432-PH-MUR and partly by the NSF Nano-scale Science and Engineering Center DMI-0327077. We thank H. Bechtel and M. C. Martin for assisting in measurements of near-infrared transmission and reflection, and S. R. J. Brueck for discussion. T.Z. acknowledges a fellowship from the Alexander von Humboldt Foundation. Multilayer deposition was performed at the Molecular Foundry, Lawrence Berkeley National Laboratory, which is supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract no. DE-AC02-05CH11231.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to X.Z. (xiang@berkeley.edu).

METHODS

In the numerical studies of the 3D fishnet metamaterial, with the exception of Figs 3c and 4d, e, we used a RCWA, which expands the electromagnetic field into 13×13 diffraction orders and matches the boundary conditions at each interface. Specifically, the numerical refractive index in Fig. 3b, the dispersion curves in Fig. 4a–c and the numerical figure of merit in Fig. 5 were calculated using a modal analysis given in ref. 22. Figures 3c and 4d, e were calculated with commercial finite-difference time-domain software (CST Microwave Studio). In the simulations, a Drude model was used for the dielectric parameters of silver, with plasma frequency $\omega_p = 9.0$ eV and scattering frequency $\gamma = 0.054$ eV. The scattering frequency is increased by a factor of three compared to that of the bulk silver³⁰ in order to account for the additional loss of surface scattering.

In the experimental setup (Fig. 2c), light from the optical parametric oscillator (Spectra-Physics) was focused onto the prism with an achromatic lens (lens 1); the second lens (lens 2) was placed at its focal position. The position of the beam at the focal distance of lens 2 (f_2) was used to calculate the angle of refraction. As a result of limited camera imaging area, only the zero-order Fourier image was recorded. To obtain the absolute angle of refraction, a window with an area equal to that of the prism was etched through the multilayer stack to serve as a reference. The window's Fourier image was measured at all wavelengths, giving a reference position corresponding to a refractive index of 1. The centres of the beam spot for both the window and prism samples were determined by fitting the intensity with a 2D Gaussian profile and the total beam shift (δ) at the position of the camera was calculated by taking the difference between the beam spot centres.

LETTERS

Chiral colloidal clusters

D. Zerrouki¹, J. Baudry¹, D. Pine², P. Chaikin² & J. Bibette¹

Chirality is an important element of biology, chemistry and physics. Once symmetry is broken and a handedness is established, biochemical pathways are set. In DNA, the double helix arises from the existence of two competing length scales, one set by the distance between monomers in the sugar backbone, and the other set by the stacking of the base pairs¹. Here we use a colloidal system to explore a simple forcing route to chiral structures. To do so we have designed magnetic colloids that, depending on both their shape and induced magnetization, self-assemble with controlled helicity. We model the two length scales with asymmetric colloidal dumbbells linked by a magnetic belt at their waist. In the presence of a magnetic field the belts assemble into a chain and the steric constraints imposed by the asymmetric spheres force the chain to coil. We show that if the size ratio between the spheres is large enough, a single helicity is adopted, right or left. The realization of chiral colloidal clusters opens up a new link between colloidal science and chemistry. These colloidal clusters may also find use as mesopolymers, as optical and light-activated structures², and as models for enantiomeric separation.

To advance the use of colloids in nanotechnology as a route for the fabrication of large-scale complex and useful structures, we need to give the building blocks the functionality and variability found in chemistry, allowing the synthesis of polymers and chiral structures rather than the crystalline structures or aggregates traditionally observed in colloids^{3–7}. Using colloids to mimic atoms, as building blocks for molecules or macromolecules at a colloidal scale, would also offer insight in designing and rationalizing functionality in chemistry. Steric hindrance between substituents is broadly accepted to be one of the main driving forces directing the formation of a particular configurational isomer⁸. At a colloidal scale, these steric interactions are much easier to control without the confounding influence of hydrogen bonding, van der Waals and electrostatic forces. But although isotropic particles suffice to form amorphous aggregates, crystals or geometrically confined structures⁹, the creation of colloidal molecules requires that the building blocks assemble in a more precise way in order to mimic bond angles. An effective processing route is to apply an external magnetic field to colloids that contain super-paramagnetic material. This way, very long polymeric assemblies can be created¹⁰. This strategy leads to very rapid assembly, but it has been limited to one-dimensional and uniform structures, thus excluding the formation of configurational isomers.

Here we introduce new magnetic colloids that can rapidly self-assemble into a variety of isomeric forms. Depending on the shape of the colloidal blocks relative to their spontaneous direction of magnetization, these colloids assemble into structures that must fulfil both steric and magnetic constraints. This induces a fixed repeated isomeric configuration or tacticity. We present two types of colloidal assemblies which led us to a third structure—a helix. The first grows in-plane with a bond angle of 60° producing a perfect alternation of colloidal block positions, a syndiotactic arrangement. The second

grows with a repeating in-plane/out-of-plane structure with a 90° rotation angle between two blocks. The third is our helical structure.

Individual colloidal blocks are produced using emulsion confinement¹¹. We start with droplets containing a binary colloidal mixture comprising both an oil-based ferrofluid and hydrophobic micrometre-sized silica particles. By adjusting the hydrophobicity of the silica particles and the quantity of iron oxide nanoparticles, we can design a specific magnetic patterning of the silica spheres during solvent evaporation. Under different conditions, we obtain four distinct types of colloidal structures after oil evaporation. Scanning electron microscope pictures of these new colloids are shown in Fig. 1. The detailed experimental protocol is described in the Methods. When no silica particles are trapped, we obtain conventional spherical magnetic colloidal particles (Fig. 1a). When only one silica particle is trapped and when both its initial size and ferrofluid content are appropriately set, we obtain silica particles that possess a single characteristic magnetic ‘cap’ (Fig. 1b). This magnetic cap consists of densely packed Fe₂O₃ nanoparticles about 10 nm in diameter. The de-wetting of the ferrofluid on the curved hydrophobic silica surface during the extraction of the oil drives its shape. At the very end of the evaporation, the nanoparticles irreversibly stick together and adhere to the silica surface, leading to the final structure. When two particles are trapped within one emulsion droplet, we obtain doublets of silica particles with a solid magnetic ring located around the contact point between the two silica spheres. This ring is shaped both by the silica contour and by capillary forces confining the ferrofluid during oil evaporation. We refer to these magnetic structures as ‘rings’ and the binary particles as ‘dumbbells’. We have made these dumbbells with either the same or differently sized silica particles, as shown in Fig. 1c and d.

Paramagnetic spherical particles as in Fig. 1a do not individually orient in a magnetic field, but two particles will attract and align in the field direction owing to shape anisotropy (part a, Fig. 2A, B). The magnetic caps and rings on the silica colloids are not magnetically isotropic and their shape anisotropy causes them to align individually in an applied field with their circular base parallel to the field¹². Two such particles will attract but their steric interaction will cause the pair not only to align with the field but to adopt specific configurations relative to each other. The capped particles have a magnetic dipole interaction which is maximized when the circular bases of the caps are parallel to each other, parallel to the field and as close together as possible. The capped particles can align with their caps on the same or opposite sides of the particles. In the latter arrangement the caps are closer (part b, Fig. 2A, B). On the symmetric dumbbells the rings form a small ‘belt’ around the waist of the touching spheres ($r_{\text{ring}} \ll r_{\text{sphere}}$). The attractive $1/r^3$ interaction is largest when the belts are as close as possible along the field. This occurs when the dumbbells are perpendicular to the magnetic field and to each other (part c, Fig. 2A, B). Two asymmetric dumbbells also adopt a ‘closest belt’ approach with all spheres touching (part d, Fig. 2A, B).

¹Laboratoire Colloïdes et Matériaux Divisés, ESPCI, UPMC, CNRS, ParisTech, 10 rue Vauquelin, 75005 Paris, France. ²Center for Soft Matter Research, Department of Physics, New York University, 4 Washington Place, New York, New York 10003, USA.

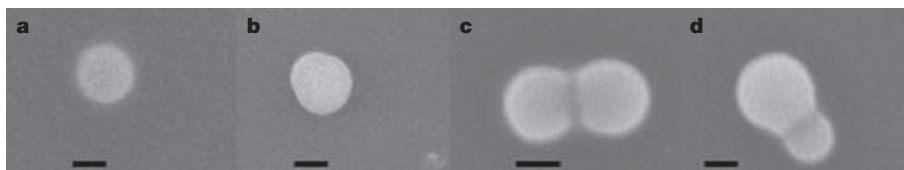


Figure 1 | Electron microscopy images of particles showing magnetic patterning. Typical scanning electron microscope pictures of dried alcano-based ferrofluid droplets. **a**, Ferrofluid only; the ferrofluid nanoparticles collapse into a spherical colloidal particle. **b**, Ferrofluid with one silica particle of diameter 1 μm ; the particle is decorated by a magnetic pattern consisting of nanoparticles of Fe_2O_3 . The cap is clearly visible when the

particle is oriented as shown here. **c**, Ferrofluid with two identical silica particles of diameter 1 μm . In this case the magnetic part consists of a ring in between the two silica spheres. **d**, Ferrofluid with different particle sizes: the silica particles are 1 and 2 μm in diameter. Owing to the purification procedure we systematically find about 10% of other types of clusters in the final samples. Scale bars, 1 μm .

In Fig. 3 we show our four different colloidal samples under a magnetic field. In all cases the attractive induced dipole–dipole interaction is sufficient to drive the self-assembly. In Fig. 3a, isotropic spherical particles form linear chains, as expected. In Fig. 3b, following the steric repulsion that we have seen in pairs in part b of Fig. 2A, B, ‘caps’ order alternately up and down along the chain direction leading to isomeric tacticity at the colloidal scale. Indeed, this assembly is reminiscent of a syndiotactic arrangement observed in linear polymeric chains, for which steric hindrance of substituents drives the regular alternation of particles¹³. For sufficiently thin caps the structure adopts a planar configuration with a bond angle of about 60° and a torsional and bond rigidity that depends on field strength. Because of the colloidal scale, the dynamics of this mode can be directly recorded in real time using a microscope and video camera (see Supplementary Movie 1). Symmetric dumbbells (Fig. 3c) self-assemble with the neighbouring dumbbells perpendicular to each other and to the field direction. Just as for chains of caps, at lower fields, thermally induced torsional modes can be excited and directly recorded in real time with a microscope (see Supplementary Movie 2).

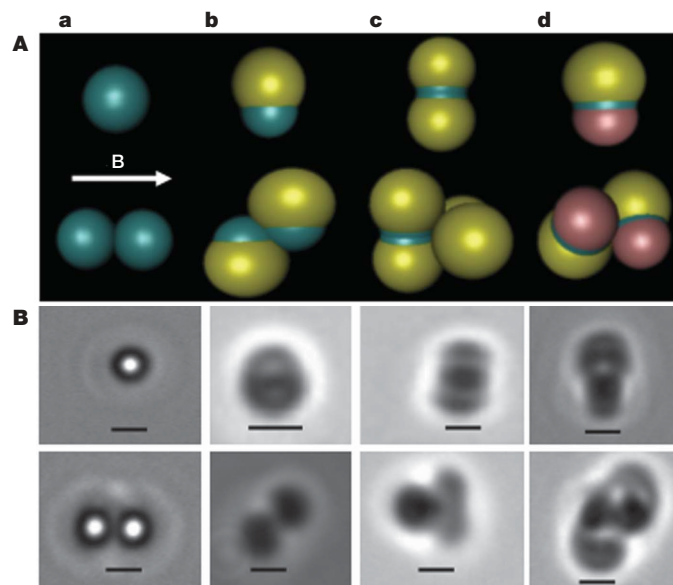


Figure 2 | Competition between magnetic interactions and steric hindrance. **A**, Schematic diagrams of the magnetic and steric pair interaction between colloidal blocks of types **a** to **d**. **a**, One isotropic paramagnetic particle has no preferential orientation in the field **B**; two of them attract to form a doublet oriented in the field direction. **b**, One capped particle orients such that the cap base is parallel to the field; two of them adjust their position so that the bases align and touch each other, forcing this up-and-down configuration. **c**, One symmetric dumbbell orients in the field so that the ring plane becomes parallel to the field; two of them must then rotate by 90° relative to each other such that the ring planes are as close as possible while retaining the right orientation. **d**, The same argument holds for a pair of asymmetric dumbbells. **B**, Microscopic pictures of single and paired building blocks of each type, following the same ordering as in **A**. Magnetic field, 10 mT; scale bars, 1 μm .

We now come to the point of emergence of chirality. It is known that monodisperse spheres interacting through excluded volume interactions and confined within a proper geometry can pack with a helical structure⁹. Here we aim at demonstrating what could be the simplest building blocks for self-assembling mesoscopic helical structures in the absence of confinement constraints. First note that in contrast to the symmetric dumbbells of Fig. 3c, a pair of magnetically linked asymmetric dumbbells, the case shown in Fig. 3d is in principle already chiral: mirror images do not superimpose, owing to the presence of the rings. Indeed, if we look at the tetrahedron formed by the spheres alone, there are two orthogonal mirror planes, but the reflection symmetry is broken by the magnetic rings (part a, Fig. 4A). A reflection plane can be restored if a third dumbbell is added symmetrically (part b, Fig. 4A). However, a more favourable magnetic interaction, with a more collinear alignment of the rings, occurs if we flip the last dumbbell (part c, Fig. 4A). The combination of a more collinear magnetic interaction plus steric repulsion favours the helical structure when the ratio of the spheres is within a critical range of values, which our calculations show to be from about 2 to 3. The calculations include the dipole–dipole interactions between rings as well as the hard sphere interactions between different spheres, but neglect the role of temperature, which must decrease the lower limit. Our first choice of a ratio of 2 for our dumbbells was based on this calculation. The movie frames in Fig. 4B show how a single dumbbell approaches a chiral cluster, then rotates and links

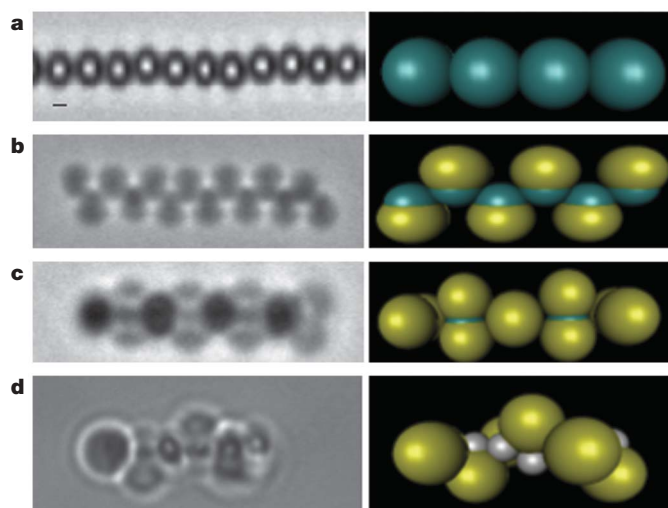


Figure 3 | Self-assembly of blocks into chains under a magnetic field. **a–d**, Optical microscope images and their corresponding schematic representations. **a**, Isotropic particles self-assemble into linear chains. **b**, Caps self-assemble by ordering alternately up and down along the chain direction. **c**, To satisfy both steric and magnetic constraints, symmetric dumbbells must rotate by 90° relative to their neighbours along the chain and field direction. **d**, Asymmetric dumbbells self-assemble into a helical structure, as a result of the steric hindrance induced by the size difference. Magnetic field, 10 mT; scale bars, 1 μm .

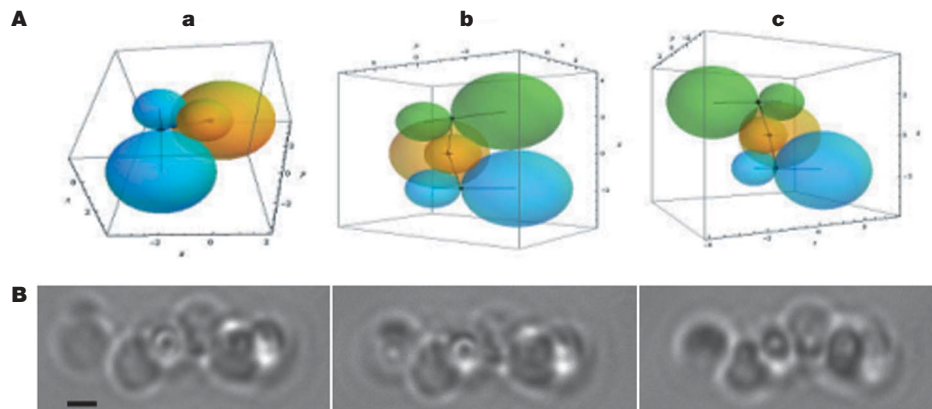


Figure 4 | Configurations and kinetic pathway for dumbbell assembly. **A**, Schematics showing a pair of dumbbells (**a**), and the two ways in which a third dumbbell could be added to that pair (**b** and **c**). **B**, Optical microscopic

images showing the kinetic pathway of a dumbbell that diffuses and assembles onto a five-dumbbell helix. Scale bar, 1 μm .

into the correct helical configuration. Of course, the kinetic pathways are complex and require further investigation, including the role of entropy and hydrodynamics. Nevertheless, we emphasize that in spite of the remaining questions about kinetic pathways and the related distributions of chiral isomers, the configurational energy is lower for helical stacking when the ratio of sphere sizes is within the critical range 2 to 3, and provides insight into the helix formation we observe.

These experiments demonstrate that colloids can build interesting conformational isomers. By controlling the shape and interactions between colloidal building blocks, a variety of structures can be designed. Here we have restricted our efforts to the growth of syndiotactic and chiral structures. We believe, however, that this work opens new routes to produce and self-assemble a wide variety of structures based on the same principles. Moreover, we believe this work could offer some guidance to the assembly of nanoparticles, which has interesting optical applications.

METHODS SUMMARY

Hydrophilic silica colloids are synthesized by hydrolysis of tetraethylorthosilicate in aqueous solution containing ethanol and ammonia, using a modified seeded growth procedure originally described by Stober and Fink¹⁴. The silica particles are surface-coated with octyldimethylchlorosilane to redisperse them into octane-based ferrofluid. This mixture is then gently emulsified into an aqueous phase containing Pluronic F127 surfactant and polyvinylpyrrolidone. The emulsion is further transformed under controlled shear into a quasi-monodisperse emulsion with globules of diameter $\sim 3\ \mu\text{m}$ (ref. 15). The final diluted emulsion is treated with a Rotary Evaporator at 80°C under reduced pressure for 2 h to evaporate the octane that remains in the globules. Cluster separation is obtained with a column that has a density gradient established by stacking solutions of various sucrose concentrations in deionized water. After 24 h of sedimentation, bands are observed and well separated such that they can be pumped out with a syringe; each fraction is then magnetically sedimented and washed several times. The colloids are examined with an optical microscope equipped with a controllable magnetic field induced by coils. For each type of building block a 50- μm -thick cell is filled with a solution with a 0.01% volume fraction of colloidal blocks and allowed to sediment for 1 min. For blocks as in parts a, b and c of Fig. 2A, a homogeneous magnetic field of 10 mT is applied. Blocks self-assemble instantly. For blocks as in part d (asymmetric dumbbells), a field of 10 mT is gradually applied over 1 min. This ensures that a larger

proportion of dumbbells assemble within the lower-magnetic-energy configuration in order to nucleate helices. Colloids in their dried state are imaged with a scanning electron microscope in secondary electron mode after sputter-coating with gold.

Received 28 February; accepted 4 July 2008.

- Watson, J. D. & Crick, F. H. C. A structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
- Dreyfus, R. *et al.* Microscopic artificial swimmers. *Nature* **437**, 862–865 (2005).
- Leunissen, M. E. *et al.* Ionic colloidal crystals of oppositely charged particles. *Nature* **437**, 235–240 (2005).
- Savage, J. R., Blair, D. W., Levine, A. J., Guyer, R. A. & Dinsmore, A. D. Imaging the sublimation dynamics of colloidal crystallites. *Science* **314**, 795–798 (2006).
- Stradner, A. *et al.* Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* **432**, 492–495 (2004).
- Gasser, U., Weeks, E. R., Schofield, A., Pusey, P. N. & Weitz, D. A. Real-space imaging of nucleation and growth in colloidal crystallization. *Science* **292**, 258–262 (2001).
- Bibette, J., Mason, T. G., Gang, H. & Weitz, D. A. Kinetically induced ordering in gelation of emulsions. *Phys. Rev. Lett.* **69**, 981–984 (1992).
- Clayden, J., Greeves, N., Warren, S. & Wothers, P. *Organic Chemistry* 1st edn (Oxford Univ. Press, 2000).
- Yin, Y. & Xia, Y. Self-assembly of spherical colloids into helical chains with well-controlled handedness. *J. Am. Chem. Soc.* **125**, 2048–2049 (2003).
- Goubault, C., Leal-Calderon, F., Viovy, J. L. & Bibette, J. Self-assembled magnetic nanowires made irreversible by polymer bridging. *Langmuir* **21**, 3725–3729 (2005).
- Manoharan, V. N., Elsesser, M. T. & Pine, D. J. Dense packing and symmetry in small clusters of microspheres. *Science* **301**, 483–487 (2003).
- Tandon, S., Beleggia, M., Zhu, Y. & De Graef, M. On the computation of the demagnetization tensor for uniformly magnetized particles of arbitrary shape. Part I: Analytical approach. *J. Magn. Magn. Mater.* **271**, 27–38 (2004).
- Young, R. J. & Lovell, P. A. *Introduction to Polymers* 2nd edn (Chapman and Hall, 1991).
- Stöber, W., Fink, A. & Bohn, E. Controlled growth of monodisperse silica spheres in the micron size range. *J. Colloid Interface Sci.* **26**, 62–66 (1968).
- Zerrouki, D. *et al.* Preparation of doublet, triangular, and tetrahedral colloidal clusters by controlled emulsification. *Langmuir* **22**, 57–62 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank R. Dreyfus for discussions.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.B. (jerome.bibette@espci.fr).

Prolonged suppression of ecosystem carbon dioxide uptake after an anomalously warm year

John A. Arnone III¹, Paul S. J. Verburg¹, Dale W. Johnson², Jessica D. Larsen¹, Richard L. Jasoni¹, Annmarie J. Lucchesi¹, Candace M. Batts¹, Christopher von Nagy¹, William G. Coulombe¹, David E. Schorran¹, Paul E. Buck¹, Bobby H. Braswell³, James S. Coleman⁴, Rebecca A. Sherry⁵, Linda L. Wallace⁵, Yiqi Luo⁵ & David S. Schimel⁶

Terrestrial ecosystems control carbon dioxide fluxes to and from the atmosphere^{1,2} through photosynthesis and respiration, a balance between net primary productivity and heterotrophic respiration, that determines whether an ecosystem is sequestering carbon or releasing it to the atmosphere. Global^{1,3–5} and site-specific⁶ data sets have demonstrated that climate and climate variability influence biogeochemical processes that determine net ecosystem carbon dioxide exchange (NEE) at multiple timescales. Experimental data necessary to quantify impacts of a single climate variable, such as temperature anomalies, on NEE and carbon sequestration of ecosystems at interannual timescales have been lacking. This derives from an inability of field studies to avoid the confounding effects of natural intra-annual and interannual variability in temperature and precipitation. Here we present results from a four-year study using replicate 12,000-kg intact tallgrass prairie monoliths located in four 184-m³ enclosed lysimeters⁷. We exposed 6 of 12 monoliths to an anomalously warm year in the second year of the study⁸ and continuously quantified rates of ecosystem processes, including NEE. We find that warming decreases NEE in both the extreme year and the following year by inducing drought that suppresses net primary productivity in the extreme year and by stimulating heterotrophic respiration of soil biota in the subsequent year. Our data indicate that two years are required for NEE in the previously warmed experimental ecosystems to recover to levels measured in the control ecosystems. This time lag caused net ecosystem carbon sequestration in previously warmed ecosystems to be decreased threefold over the study period, compared with control ecosystems. Our findings suggest that more frequent anomalously warm years⁹, a possible consequence of increasing anthropogenic carbon dioxide levels¹⁰, may lead to a sustained decrease in carbon dioxide uptake by terrestrial ecosystems.

Ecosystem biogeochemical processes that modulate CO₂ exchange between land and atmosphere respond to climate variability at different timescales. Well known and characterized are the effects of seasonal changes in terrestrial net primary productivity (NPP) on NEE (the instantaneous net ecosystem CO₂ flux) that cause intra-annual fluctuations in global atmospheric CO₂ levels¹. Less well quantitatively understood, however, are potential delayed or lagged effects of interannual climate variability—particularly climatically anomalous years—on net ecosystem productivity (NEP, the annual sum of NEE) that in turn help to determine interannual rates of change of atmospheric CO₂ concentration (refs 11, 12). This inadequate quantitative understanding of lagged responses to climatically

anomalous years is of particular concern because, as a consequence of increasing anthropogenic CO₂ concentrations (ref. 10), the frequency and intensity of extreme years are increasing⁹.

Although statistical and process modelling of global and site-specific temperature, precipitation, [CO₂] and NEE data at intra-annual and interannual timescales suggest a strong temporal correlation between climate variability, vegetation and growth rate of atmospheric CO₂ concentration (refs 1, 11, 12), interannual variability in both temperature and precipitation makes a quantitative examination of underlying ecosystem control mechanisms by using field data difficult. The EcoCELL facility at the Desert Research Institute uniquely enables: (1) tests of the effects of variability in a single climate factor at interannual timescales while simulating natural diel and seasonal climate variation; (2) continuous monitoring of NEE and ecosystem processes that determine NEE for multiple years; and (3) the study of intact soil monoliths with their native plant communities^{7,8}.

Key objectives of the present study were to quantify, first, how exposure to an anomalously warm year affects ecosystem processes that determine NEE and NEP, and thus annual net C sequestration, at intra-annual and interannual timescales; second, how intra-annual responses and feedbacks shape interannual responses; and third, the environmental factors and feedbacks that affect these processes. We used tallgrass prairie as a model ecosystem (see Methods).

During the course of the study, the intact natural ecosystems used in the EcoCELLs behaved largely in a fashion representative of natural tallgrass prairie ecosystems in the field. Plant canopies developed and senesced in a pattern (Fig. 1e) that followed seasonal changes in air temperature, daytime vapour pressure deficit (VPD) and rainfall (Fig. 1a, b, d). NEE and evapotranspiration closely followed patterns in plant canopy greenness (Fig. 1c, e, f).

In the treatment year, ecosystems exposed to temperatures 4 °C higher responded with a 7–10-day earlier start of the growing season, resulting in a more rapid rise in NEE and evapotranspiration in early spring (mid March to early April 2003; Fig. 1c, f). Mean daytime VPDs in warmed EcoCELLs immediately doubled when temperatures were raised, and this increased evapotranspiration in early spring 2003 (Fig. 1b, c). By the end of April 2003, NEE peaked almost 35 days earlier in warmed temperatures than in control temperatures (Fig. 1f). This coincided with an increase in mean daytime VPD, a disappearance of the treatment effect on evapotranspiration, a decline in soil water content (SWC) in surface horizons, and a slowing of plant canopy development (Fig. 1b–d and Supplementary

¹Desert Research Institute, Reno, Nevada 89512, USA. ²Department of Natural Resources and Environmental Science, University of Nevada, Reno, Nevada 89557, USA. ³Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Durham, New Hampshire 03824, USA. ⁴Office of Research and Department of Ecology and Evolutionary Biology, Rice University, Houston, Texas 77251, USA. ⁵Department of Botany and Microbiology, University of Oklahoma, Norman, Oklahoma 73019, USA. ⁶National Center for Atmospheric Research, Boulder, Colorado 80305, USA.

Fig. 3). Although higher evapotranspiration rates in warmed ecosystems in early spring led to slightly lower SWC than in control ecosystems, significant decreases in SWC became apparent only in July 2003 (Fig. 1d). Daytime VPDs in warmed EcoCELLs reached levels high enough (more than 10 mbar) to decrease leaf stomatal conductance (g_s)^{13–15} by late March 2003, more than two months earlier than in control EcoCELLs. VPDs in warmed EcoCELLs remained above 10 mbar twice as long as they did in control EcoCELLs.

Thus, the earlier decline in NEE observed in warmed ecosystems in 2003 seems to have resulted, at least initially, from higher daytime VPDs that substantially constrained canopy-level CO_2 uptake¹⁶. The importance of VPD as a regulator of canopy conductance (g_{canopy}) has been demonstrated in a wide range of graminoid-dominated ecosystems^{8,17–19}, and decreases in g_s caused by high VPDs¹⁵ have been shown to decrease canopy⁸ and leaf-level¹⁴ CO_2 assimilation. Lower CO_2 uptake by the canopy may have led to decreased canopy

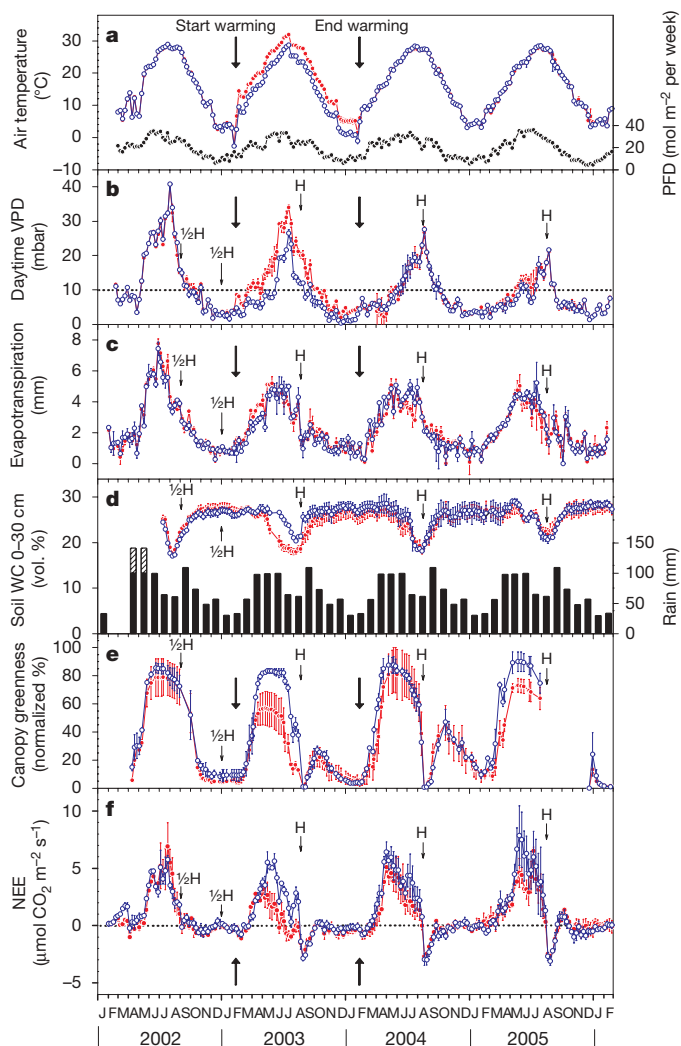


Figure 1 | Four-year time courses covering pretreatment (11 February 2002 to 10 February 2003), treatment and post-treatment years. Weekly means are shown. **a**, Air temperature and photon flux density (PFD). **b**, Daytime atmospheric VPD with line at 10 mbar denoting VPD above which g_s can be strongly decreased. **c**, Ecosystem evapotranspiration. **d**, SWC of the 0–30 cm layer; bars show applied monthly rainfall (lighter-shaded bar-tops in April and May 2002 denote additional rain applied to compensate for below-target rainfall in February and March 2002). **e**, Plant canopy greenness index. **f**, NEE based on 24 h means; positive values indicate net CO_2 uptake by the ecosystem, and negative values net CO_2 release. Aboveground plant biomass harvests are indicated by H, and half-plot harvests by ½H. Open blue circles, control ecosystems; filled red circles, treated ecosystems (warmed in 2003). Error bars indicate s.e.m. for $n = 2$ EcoCELLs.

development and duration (also see Supplementary Information) and thus to a lower NPP, which limited NEE even further in conjunction with a decrease in soil CO_2 efflux⁸. This VPD response may have been exacerbated by drier surface soils in warmed ecosystems beginning in May 2003 (Fig. 1d), potentially lowering the apparent g_s . Decreases in mean weekly NEE observed in warmed ecosystems in the treatment year (Fig. 1f) resulted mainly from large decreases in daytime NEE (Supplementary Fig. 4a) that may have been due partly to higher aboveground plant respiration immediately following the step increase in temperature (Supplementary Fig. 4b, inset, and Supplementary Information). Hence, immediate and ongoing plant physiological responses to increased temperatures and VPDs in 2003 seem to have contributed to overall declines in annual ecosystem CO_2 balance primarily by decreasing NPP. Canopy green index explained between 70% and 78% of the variability observed in mean weekly NEE (Supplementary Fig. 6), with canopy clipping in late August 2003 (and August 2004) eliminating nearly all CO_2 uptake in all ecosystems and eliminating differences in NEE between control and warmed ecosystems for the rest of the treatment year (Fig. 1f).

Persistence of lower SWC in warmed ecosystems into the first post-treatment year (2004) (Fig. 1d) may explain the two-week slower start in plant canopy green-up and growth in spring of 2004 than in the controls (Fig. 1e). Canopy green-up in these ecosystems seemed to start only after springtime rains restocked topsoil moisture to levels present in unwarmed ecosystems two weeks earlier. In 2004, VPD differences between previously warmed and unwarmed ecosystems disappeared when air temperatures in treated EcoCELLs were returned to pretreatment levels (Fig. 1b). In 2005 (the second post-treatment year), the start of canopy green-up in previously warmed ecosystems was once again delayed (Fig. 1e). In contrast to the return of normal plant canopy development in previously warmed ecosystems in 2004 after the delayed start, however, plant canopy development in 2005 in previously warmed ecosystems was decreased, but not as severely as the decrease observed in 2003. In 2005, decreased canopy size did not result in significant decreases in NEE, presumably because lower VPDs in 2005 did not constrain g_s and g_s -modulated

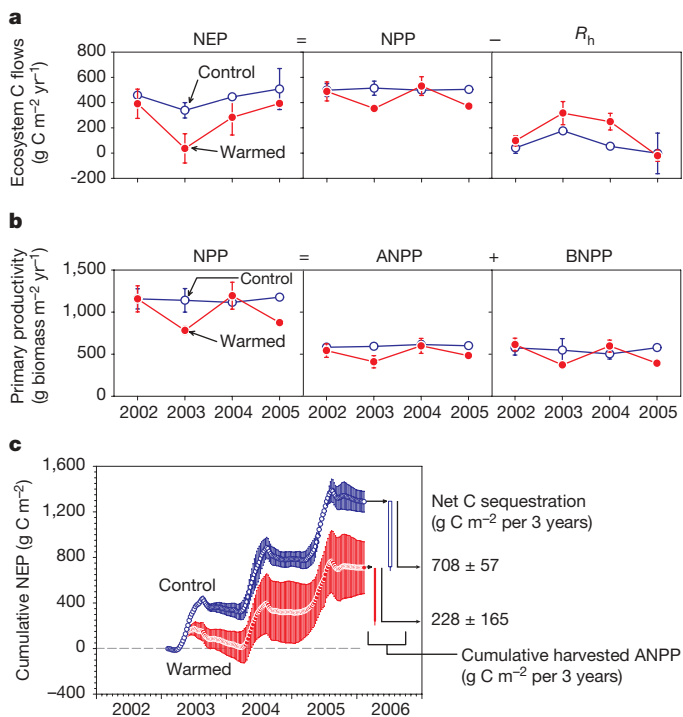


Figure 2 | Effects of an anomalously warm year on annual ecosystem C flows. **a**, NEP, NPP and R_h . **b**, Primary productivity: NPP, ANPP and below-ground NPP (BNPP). **c**, Cumulative net ecosystem productivity—NEP. Error bars indicate s.e.m. for $n = 2$ EcoCELLs.

leaf photosynthesis as much as higher VPDs in 2003 seem to have done.

Aggregated effects of the warm year included a 63% decrease in NEP in 2003, relative to the controls, followed by only a partial recovery in NEP in the first post-treatment year (2004; Figs 2a and 3) and complete recovery in the second post-treatment year (2005). Large decreases in NPP in 2003 (Fig. 2a, b) clearly contributed to the decrease in NEP in that year (see also Supplementary Information). Full recovery of NPP in 2004 enabled a partial recovery of NEP in 2004. Full recovery of NEP in 2005 occurred despite a two-year lagged decrease in NPP. However, this decrease in NPP may, in turn, have resulted in decreased C supplies to soil heterotrophs, thus also decreasing heterotrophic respiration (R_h) and enabling recovery of NEP.

Our results indicate that a lack of complete NEP recovery in 2004 was caused by a lagged warming-induced stimulation of R_h (calculated as NPP minus NEP; Fig. 2a). Conversely, full recovery of NEP in 2005 in previously warmed ecosystems was possible only because R_h of these ecosystems returned to levels measured in the pretreatment year in ecosystems assigned to the warming treatment. The absence of significant stimulation of R_h in warmed ecosystems in 2003 may have been caused by large decreases in the SWC of the uppermost soil horizons (Fig. 1d), suppressing soil microbial activity. Stimulation of R_h in 2004 in previously warmed ecosystems may have resulted from a recovery of soil moisture levels (and possibly plant-available soil N levels) that enabled a breakdown of undecomposed rhizodeposits produced in 2003 along with a breakdown of fresh rhizodeposits produced in 2004 by fully recovered plant canopies (Fig. 1e), photosynthate supplies and NPP (Fig. 2a, b). Causes of the unexpected recurring decrease in NPP and in plant canopy development in the second post-treatment year in previously warmed ecosystems are unclear but may have included a decrease in labile N caused by the previous year's post-warming rebounding aboveground NPP (ANPP) that was harvested in the late summer of 2004. Regardless of the cause, the NPP response observed in 2005 and the response showing lower R_h in 2005 in previously warmed ecosystems demonstrate the potential for longer-term and unexpected lagged responses that may affect the CO_2 balance of an ecosystem.

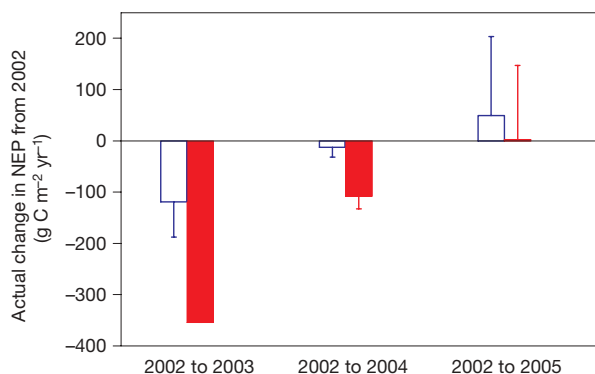


Figure 3 | Actual change in ecosystem annual NEP for control and warmed ecosystems from 2002 to 2003, 2002 to 2004, and 2002 to 2005. Results are means and s.e.m. ($n = 2$); the error bar for 'warmed 2002 to 2003' was $\pm 0.11 \text{ g C m}^{-2} \text{ yr}^{-1}$ and is not visible at the y-axis scale used in the figure. This figure also shows a strong warming-induced decrease in annual NEP in warmed ecosystems during the treatment year, only partial recovery in 2004, and complete recovery in 2005. Open blue columns, control ecosystems for 2002, 2004 and 2005; filled red columns, treated ecosystems (warmed in 2003) for 2002, 2004 and 2005. The graph shows that the lagged decrease in NEP that occurred in the year after the anomalously warm year was about one-third as large as the synchronous warming-induced decrease that occurred in the warm year (seen by comparing the first two red bars with each other). When accounting for changes in NEP in the control ecosystems in both years, however, the true effect was 40% of the immediate effect.

Data from this study and other field studies^{20–23} suggest that warming affects NEE and thus NEP in the year of the temperature anomaly primarily through hydrological feedbacks on plant canopy physiology (g , and CO_2 assimilation), development, size (greenness) and duration (Fig. 1b, e, f, and Supplementary Figs 6, 7 and 8; see also Supplementary Information). In fact, declines in ANPP observed in tallgrass prairie in the field in years with low precipitation²⁴ were similar to those that we observed in 2003 in response to the warm year. The responses we observed in subsequent years seem to have been due to more complex interacting feedbacks, which may include lagged decomposition of plant residues, carryover effects of the previous year's water deficits, and feedbacks through nutrient cycles. The collective result of synchronous and lagged declines in NEP observed during the four years of our study was a persistent decrease in annual net ecosystem C sequestration (NEP minus C removed in annual ANPP harvests that would decompose and release CO_2) that led to a threefold decrease by the end of the fourth year of the experiment (2005; Fig. 2c). CO_2 fluxes measured in our study were realistic and comparable to those measured in tallgrass prairie in the field²⁵.

Taken together, results from this multiyear experiment demonstrate, first, that the response of ecosystems to climate variability seems to be more complex than can easily be inferred from traditional experimental and observational approaches; second, that an increase in frequency and intensity of anomalously warm years may decrease the ability of terrestrial ecosystems to absorb CO_2 and store carbon^{22,23}—more than would be expected on the basis of previous experimental approaches—and third, that the lagged effects of climate anomalies on NEP (less than observed during the year of the anomaly) provide a basis for validating timescales and mechanisms used in models²⁶, as well as projecting actual ecosystem C sequestration and calculating carbon credits across multiple years.

METHODS SUMMARY

We installed three large monoliths (2.44 m long, 1.22 m wide and 1.80 m deep) of intact soil and their actual plant communities excavated from a C_4 -dominated tallgrass prairie in central Oklahoma, USA (Supplementary Fig. 1; also see Supplementary Information) into each of four 184-m^3 (7.5 m long, 5.5 m wide and 4.5 m high) daylight EcoCELLs^{7,8} that served as both individual controlled-environment chambers and large differential open-flow whole-ecosystem gas exchange cuvettes enabling the continuous measurement of NEE. Air temperatures in the EcoCELLs in all four years of the experiment (2002–2005) were programmed to follow natural diel and seasonal oscillations based on data collected at the excavation site from 1993 to 2000 with the temperature in two of the EcoCELLs increased by 4°C (Fig. 1a) in year 2 (the anomalously warm year) and then returned to the pretreatment regime in year 3. Annual rainfall was held constant in all four years (980 mm) and applied at the average monthly amount and natural mean frequency for each month (Fig. 1d). We monitored NEE continuously by measuring the difference between the $[\text{CO}_2]$ in the air entering each EcoCELL and the $[\text{CO}_2]$ of the mixed air inside each EcoCELL and multiplying this difference by the mass flow of air passing through the EcoCELL (20 mol s^{-1} ; see Supplementary Methods). This value was then divided by the total land area of the three monoliths to yield NEE in units of $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$. The annual NEP for each EcoCELL was calculated as the sum of all NEE values measured in an EcoCELL during the year. We measured annual NPP, ANPP and below-ground NPP (BNPP). We also continuously monitored evapotranspiration (daily), SWC (hourly) and plant canopy green index (weekly). Annual R_h was calculated for each EcoCELL as $R_h = \text{NPP minus NEP}$. Ecosystem C sequestration was calculated as NEP minus harvested ANPP-C. We analysed treatment effects on interannual and intra-annual data sets for all parameters primarily by using repeated-measures analysis of variance with each EcoCELL as a replicate ($n = 2$; Supplementary Table 1; also see Supplementary Methods).

Received 7 November 2007; accepted 29 July 2008.

- Keeling, C. D., Chin, J. F. S. & Whorf, T. P. Increased activity of northern vegetation inferred from atmospheric CO_2 measurements. *Nature* **382**, 146–149 (1996).
- Schimel, D. S. Terrestrial ecosystems and the carbon-cycle. *Glob. Change Biol.* **1**, 77–91 (1995).
- Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **386**, 698–702 (1997).

4. Randerson, J. T., Thompson, M. V., Conway, T. J., Fung, I. Y. & Field, C. B. The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide. *Glob. Biogeochem. Cycles* **11**, 535–560 (1997).
5. Vukicevic, T., Braswell, B. H. & Schimel, D. A diagnostic study of temperature controls on global terrestrial carbon exchange. *Tellus* **53B**, 150–170 (2001).
6. Law, B. E. *et al.* Environmental controls over carbon dioxide and water vapor exchange of terrestrial vegetation. *Agric. Forest Meteorol.* **113**, 97–120 (2002).
7. Griffin, K. L. *et al.* EcoCELLs: Tools for mesocosm scale measurements of gas exchange. *Plant Cell Environ.* **19**, 1210–1221 (1996).
8. Verburg, P. S. J., Larsen, J. D., Johnson, D. W., Schorran, D. E. & Arnone, J. A. III. Impacts of an anomalously warm year on soil CO₂ fluxes in experimentally manipulated tallgrass prairie ecosystems. *Glob. Change Biol.* **11**, 1720–1732 (2005).
9. Easterling, D. R. *et al.* Climate extremes: Observations, modeling, and impacts. *Science* **289**, 2068–2074 (2000).
10. Intergovernmental Panel on Climate Change. *Climate Change 2007: The Physical Science Basis: Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2007).
11. Braswell, B. H., Schimel, D. S., Linder, E. & Moore, B. The response of global terrestrial ecosystems to interannual temperature variability. *Science* **278**, 870–872 (1997).
12. Tian, H. Q. *et al.* Effect of interannual climate variability on carbon storage in Amazonian ecosystems. *Nature* **396**, 664–667 (1998).
13. Körner, C. in *Ecology of Photosynthesis* (eds Schulze, E. D. & Caldwell, M. M.) 463–490 (Springer, 1994).
14. Bunce, J. A. Low humidity effects on photosynthesis in single leaves of C-4 plants. *Oecologia* **54**, 233–235 (1982).
15. Oren, R. *et al.* Survey and synthesis of intra- and interspecific variation in stomatal sensitivity to vapour pressure deficit. *Plant Cell Environ.* **22**, 1515–1526 (1999).
16. Wever, L., Flanagan, L. B. & Carlson, P. J. Seasonal and interannual variation in evapotranspiration, energy balance and surface conductance in a northern temperate grassland. *Agric. For. Meteorol.* **112**, 31–49 (2002).
17. Grace, J., Lloyd, J., Miranda, A. C., Miranda, H. & Gash, J. H. C. Fluxes of carbon dioxide and water vapour over a C-4 pasture in south-western Amazonia (Brazil). *Aust. J. Plant Physiol.* **25**, 519–530 (1998).
18. Vourlitis, G. L. & Oechel, W. C. Eddy covariance measurements of CO₂ and energy fluxes of an Alaskan tussock tundra ecosystem. *Ecology* **80**, 686–701 (1999).
19. Novick, K. A. *et al.* Carbon dioxide and water vapor exchange in a warm temperate grassland. *Oecologia* **138**, 259–274 (2004).
20. Saleska, S. R., Harte, J. & Torn, M. S. The effect of experimental ecosystem warming on CO₂ fluxes in a montane meadow. *Glob. Change Biol.* **5**, 125–141 (1999).
21. Rustad, L. E. *et al.* A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia* **126**, 543–562 (2001).
22. Ciais, P. *et al.* Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature* **437**, 529–533 (2005).
23. Reichstein, M. *et al.* Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis. *Glob. Change Biol.* **13**, 634–651 (2007).
24. Knapp, A. K., Briggs, J. M. & Koelliker, J. K. Frequency and extent of water limitation to primary production in a mesic temperate grassland. *Ecosystems (N. Y., Print)* **4**, 19–28 (2001).
25. Suyker, A. E., Verma, S. B. & Burba, G. G. Interannual variability in net CO₂ exchange of a native tallgrass prairie. *Glob. Change Biol.* **9**, 255–265 (2003).
26. Schimel, D. S., Churkina, G., Braswell, B. H. & Trenbath, J. in *A History of Atmospheric CO₂ and its Effects on Plants, Animals and Ecosystems* (eds Baldwin, I. T. *et al.*) 350–368 (Springer, 2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank E. Kessler and the University of Oklahoma for the use of the Kessler Farm Field Laboratory; R. Kreidberg for editorial assistance; and J. Amthor, M. Bahn, C. Körner, D. Obrist and G. Wohlfahrt, for critical review of an early draft of this manuscript. This study was funded by the US National Science Foundation's Integrated Research Challenges in Environmental Biology program, with additional support from the Desert Research Institute (DRI) for necessary upgrades to the EcoCELL facility. We are very grateful to DRI, the University of Oklahoma, and the University of Nevada–Reno teams for technical assistance, facilities support, and administrative expertise (also see Supplementary Information).

Author Contributions J.A.A. and P.S.J.V. headed up the study including leading proposal writing, project coordination, data interpretation and analysis, and manuscript preparation. P.S.J.V., R.L.J., J.D.L., D.E.S., J.A.A. and W.G.C. were directly involved in all aspects of the study on a day-to-day basis, with A.J.L., J.A.A., L.L.W. and R.A.S. focusing on plant community aspects; C.M.B. contributing to hydrological measurements; J.D.L. contributing to quantification of plant canopy dynamics; P.S.J.V. contributing particularly to measurements of soil CO₂ fluxes and soil C and N; and D.W.J. contributing to estimation of soil nutrients. D.S.S., Y.L., B.H.B., J.S.C., P.S.J.V. and J.A.A. developed the idea for the research. C.v.N. built and managed the database for the project and contributed to data analyses. P.E.B. coordinated outreach activities to local schools and brought research to the communities of Reno and Las Vegas. R.L.J., J.D.L. and P.S.J.V. worked closely with J.A.A. on data analysis and manuscript preparation. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.A.A. III (jarnone@dri.edu).

Counterintuitive carbon-to-nutrient coupling in an Arctic pelagic ecosystem

T. F. Thingstad¹, R. G. J. Bellerby^{2,3}, G. Bratbak¹, K. Y. Børsheim⁴, J. K. Egge¹, M. Heldal¹, A. Larsen¹, C. Neill², J. Nejtgaard¹, S. Norland¹, R.-A. Sandaa¹, E. F. Skjoldal¹, T. Tanaka⁵, R. Thyrrhaug¹ & B. Töpper¹

Predicting the ocean's role in the global carbon cycle requires an understanding of the stoichiometric coupling between carbon and growth-limiting elements in biogeochemical processes. A recent addition to such knowledge is that the carbon/nitrogen ratio of inorganic consumption and release of dissolved organic matter may increase in a high-CO₂ world¹. This will, however, yield a negative feedback on atmospheric CO₂ only if the extra organic material escapes mineralization within the photic zone. Here we show, in the context of an Arctic pelagic ecosystem, how the fate and effects of added degradable organic carbon depend critically on the state of the microbial food web. When bacterial growth rate was limited by mineral nutrients, extra organic carbon accumulated in the system. When bacteria were limited by organic carbon, however, addition of labile dissolved organic carbon reduced phytoplankton biomass and activity and also the rate at which total organic carbon accumulated, explained as the result of stimulated bacterial competition for mineral nutrients. This counterintuitive 'more organic carbon gives less organic carbon' effect was particularly pronounced in diatom-dominated systems where the carbon/mineral nutrient ratio in phytoplankton production was high. Our results highlight how descriptions of present and future states of the oceanic carbon cycle require detailed understanding of the stoichiometric coupling between carbon and growth-limiting mineral nutrients in both autotrophic and heterotrophic processes.

The balance between photosynthetic production and microbial consumption of organic matter in the ocean's surface layer is not well understood, as is illustrated by the present debate on the apparent net heterotrophy of the ocean², and the lack of a good understanding of the mechanism by which dissolved organic carbon (DOC) accumulates in the upper mixed layer³. On the basis of simple models, the microbial part of the food web has been suggested to have two possible steady states³, characterized by the growth rate of heterotrophic prokaryotes (henceforth, bacteria) being limited by organic carbon (C) or by mineral nutrients (hereafter referred to as nutrients), respectively. The coupling between autotrophic and heterotrophic processes would be expected to differ considerably between these two system states.

Some of the hypothesized relationships between the structure of the microbial food web and consumption of labile organic material can be illustrated using the idealized food web⁴ in Fig. 1. In a food web dominated by osmotrophs (phytoplankton and bacteria), there would be a strong competition for nutrients combined with little recycling, favouring a state where bacterial growth rate is limited by nutrients. This situation would be further enhanced if nutrient-stressed phytoplankters excrete carbon-rich, degradable, organic compounds⁵. In a later state of succession with the biomass distribution shifted 'up'

towards phagotrophs, reduced competition and increasing recycling would furnish the bacteria with more nutrients, thereby allowing increased degradation of labile organic matter and thus favouring a shift to states where bacterial growth is limited by the supply of organic carbon. The model presented in Fig. 1 further suggests a negative interaction between labile organic matter and silicate with two extreme outcomes: either silicate leads to increased immobilization of nitrogen (N) and phosphorus (P) in diatom biomass and therefore reduces the potential for bacterial consumption of organic matter, or an increase in organic matter leads to increased bacterial sequestration of nitrogen and phosphorus and therefore to reduced diatom growth. It has also been proposed that the heterotrophic components of this system, which includes bacteria⁶, protozoa⁷ and copepods⁷, are affected more adversely than phytoplankton by low temperatures; hence, there is also a potential temperature effect on the autotroph–heterotroph balance, potentially important in the present situation with the Arctic Ocean feared to be in rapid transition^{8,9}.

To study these hypothesized nutrient–organic carbon interactions, we conducted a mesocosm study at Ny Ålesund, Svalbard (79° N) with a design intended to experimentally simulate variable organic-carbon/nutrient stoichiometry in the dissolved pool. In brief, phytoplankton blooms were induced in eight mesocosm units by adding identical daily doses of inorganic nitrogen and phosphorus corresponding to the Redfield ratio. The eight units formed two four-point gradients, differing in silicate treatment but identical in their gradient of addition rates of labile DOC as glucose (0, 0.5, 1 and 3

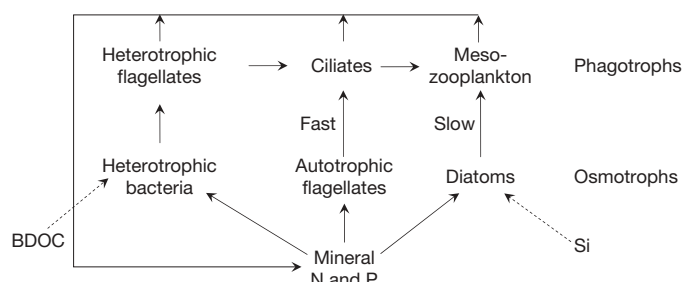


Figure 1 | Idealized microbial food web used to illustrate some of the interactions believed to be important in controlling the autotroph–heterotroph balance in the photic zone. The mesocosm experiment described here investigated the role of biodegradable DOC (BDOC) and silicate (Si, dashed arrows) in modifying the cycle of limiting nutrients (solid arrows) by potentially constraining the flows through the left (bacterial) and right (diatom) side of the network. If the numerical response to increased food supply is fast in ciliates, but slow in copepods, this will tend to prolong blooms when silicate is available.

¹Department of Biology, University of Bergen, Jahnebakken 5, PO Box 7800, 5020 Bergen, Norway. ²Bjerknes Centre for Climate Research, Allégaten 55, 5007 Bergen, Norway. ³Geophysical Institute, University of Bergen, Allégaten 70, 5007 Bergen, Norway. ⁴Institute of Marine Research, PO Box 1870, Nordnes 5817 Bergen, Norway. ⁵Laboratoire d'Océanographie et de Biogéochimie, UMR6535-CNRS, Campus de Luminy, Case 901, 13288 Marseille CEDEX09, France.

times the Redfield ratio, respectively, in terms of glucose carbon relative to the nitrogen and phosphorus additions). The two gradients were parallel in treatment up to day 4.5, when the initial concentration of free silicate became depleted. Thereafter, one gradient ($-Si$, which consisted of units 0C, 0.5C, 1C and 3C) received no experimental addition of silicate, whereas the other gradient ($+Si$, which consisted of units 0CSi, 0.5CSi, 1CSi and 3CSi) was kept silicate-replete. In an initial phase with free nutrients (phase I, up to approximately day 5), all units developed in parallel with a near-exponential net growth both in chlorophyll *a* (Fig. 2a; generation time $t_g \approx 1.7$ d) and in bacterial abundance ($t_g \approx 1.8$ d, see Supplementary Information). After nutrient depletion (phase II, days 5–12), development in the different treatments diverged (Fig. 2).

Phytoplankton community composition developed differently in the two gradients, with mixed, flagellate-dominated communities in the $-Si$ mesocosms and communities dominated by a small ($<10 \mu m$) diatom (*Thalassiosira* sp.) in $+Si$ mesocosms. Assuming

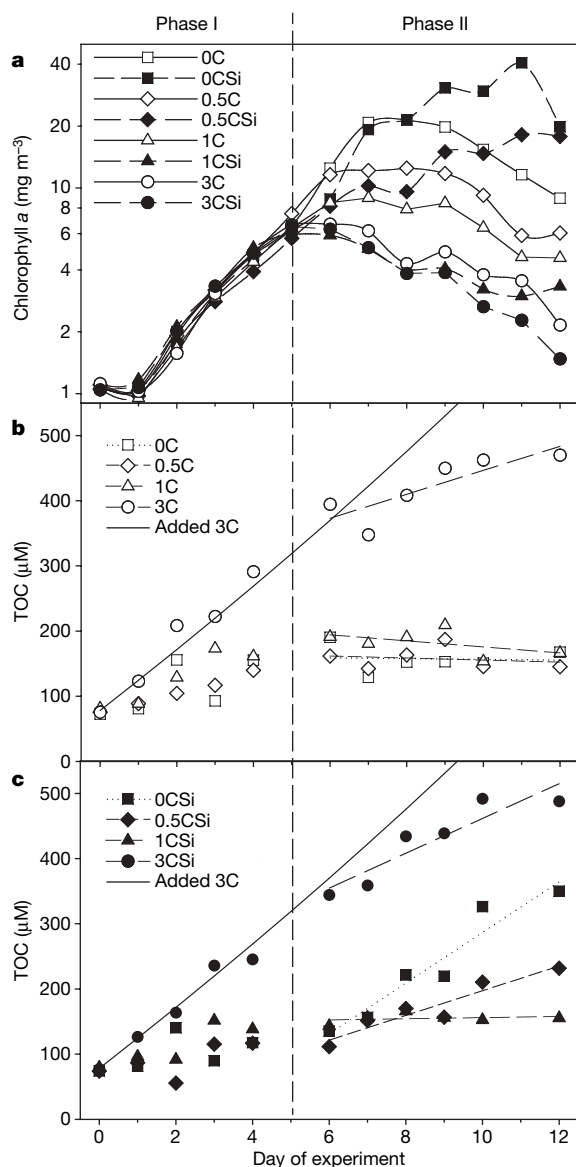


Figure 2 | Time course of bloom development. a–c, Chlorophyll *a* (a, all eight mesocosm units) and TOC (b, $-Si$ gradient; c, $+Si$ gradient). The vertical dashed line denotes the approximate time of transition to mineral nutrient depletion. The regression lines for days 5–12 used to calculate TOC accumulation rates are shown in b and c. The solid lines in b and c indicate the TOC concentration calculated assuming glucose addition as in 3C and 3CSi, combined with no biological consumption.

a slower numerical response in the diatom-to-copepod than in the corresponding autotrophic flagellate-to-ciliate succession, the model in Fig. 1 predicts larger and more prolonged blooms in the $+Si$ gradient than in the $-Si$ gradient. This is consistent with observations towards the end of phase II, when chlorophyll *a* concentrations were higher in 0CSi and 0.5CSi than in 0C and 0.5C, respectively (Fig. 2a). In both units without glucose addition (0C and 0CSi), gross primary production ($24 \mu M C$ per day and $76 \mu M C$ per day, respectively) exceeded the carbon fixation value ($\sim 18 \mu M C$ per day) obtained from converting nutrient addition rate using the Redfield ratio. Part of this excess is presumably reflecting internal recycling of nutrients, but the high value in 0CSi (4.2 times the Redfield ratio compared to nutrient addition) suggests a high carbon fixation relative to nutrient uptake in nutrient-stressed diatoms⁵.

The system response pattern shifted character at a glucose addition rate of approximately 1 times the Redfield ratio (Fig. 3). For glucose addition rates <1 times the Redfield ratio, increased glucose supply consistently gave negative trends in parameters related to primary production, particularly pronounced in the $+Si$ gradient. This is seen both as a decrease in chlorophyll *a* (Figs 2a and 3a) and in gross primary production (Fig. 3b; for statistical considerations, see Supplementary Information). The net consequence for the system's carbon balance was a reduction in accumulation rates of total organic carbon (TOC; Figs 2b, c and 3g). In the $+Si$ gradient, this 'more organic carbon gives less organic carbon' effect gave a reduction of $38 \mu M C$ per day in TOC accumulation rate, more than twice the corresponding increase in glucose addition rate between 0 times and 1 times the Redfield ratio of $18 \mu M$ glucose-C per day (Fig. 3g). Within the framework of the idealized model (Fig. 1), this can be explained as the reduction in primary production being mediated by means of nutrient competition from carbon-limited bacteria, able to sequester more nutrients when supplied with more glucose. A high carbon/nutrient ratio in diatom primary production would explain the strong enhancement of the effect in the $+Si$ gradient. In the $-Si$ gradient, there was a similar, but less pronounced, negative trend in TOC accumulation (reduction between 0 times and 1 times the Redfield ratio = $4.1 \mu M$ glucose-C per day).

For glucose addition rates beyond 1 times the Redfield ratio, effects on chlorophyll levels were minor (Fig. 3a). This is suggested to reflect a shift to nutrient-limited bacterial growth. Extra additions of organic carbon to a system where bacterial growth is already nutrient-limited (organic carbon replete) would not be expected to have major consequences for trophic interactions. The main expected effect would be an accumulation of the additional supply as non-degraded (possibly transformed) organic carbon³. From 1 times to 3 times the Redfield ratio, the observed increase in accumulation rate of TOC was $\sim 25 \mu M C$ per day in both gradients (Fig. 3g). This is smaller than, but comparable to, the corresponding increase in glucose addition rate of $\sim 36 \mu M C$ per day, consistent with this expectation.

Bacterial abundance (Fig. 3c), community respiration (Fig. 3d) and leucine incorporation (Fig. 3e) all increased along the gradients, mirroring the negative effect on phytoplankton. The overall result was a shift towards net heterotrophy between 0.5 times and 1 times the Redfield ratio in both gradients (Fig. 3f). A conversion of leucine incorporation (Fig. 3e) to nitrogen consumption assuming $1.5 \text{ kg C per mol leucine}^{10}$ and a molar nitrogen/carbon ratio of 0.27 in marine bacteria¹¹, indicates that bacterial nitrogen consumption increased from about 11% to about 120% of added nitrogen from one end of the gradients to the other, supporting the proposed switch in phytoplankton–bacteria nutrient partitioning. For organic material to accumulate in situations with carbon-limited bacteria, the material should be unavailable to bacteria. The high accumulation rates of TOC in 0CSi and 0.5CSi therefore indicates that a high proportion of the material produced autochthonously is unavailable to bacteria over the timescale studied here; this may be unavailable either because of its chemical form or because of physical protection (for example, inside diatoms). Bacterial abundance (Fig. 3c), community

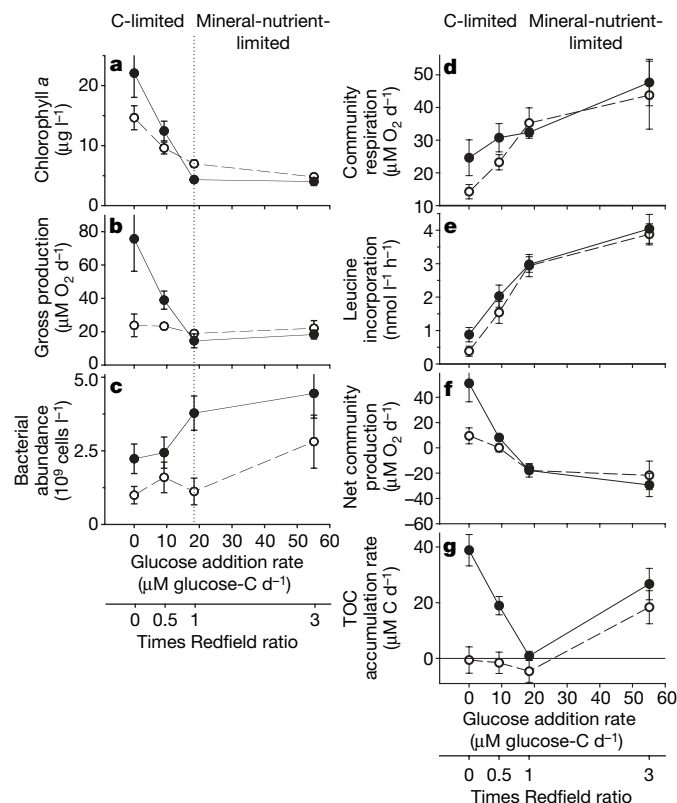


Figure 3 | Ecosystem responses to glucose. The average of one measurement per day for days 5–12 for autotrophic (a, b) and heterotrophic (c–e) food-web components, and their balance along the –Si (open circle) and +Si (filled circle) gradients (f). (a, Chlorophyll *a*; b, gross production; c, bacterial abundance; d, community respiration; e, leucine incorporation; f, net community production.) The dotted vertical line indicates shift from carbon-limited to nutrient-limited bacterial growth with qualitatively different responses in TOC accumulation rates to increasing glucose addition rate in the two regions (g). Error bars, s.e.m.; $n = 8$ in a, c and e, $n = 4$ in b, d and f. Error bars in g represent standard error of slopes of linear regressions ($n = 7$) in Figs 2b, c.

respiration (Fig. 3d) and leucine incorporation (Fig. 3e) continued to increase when glucose addition rates were increased into the region with presumed nutrient limitation of bacterial growth (higher values in 3C and 3CSi than in 1C and 1CSi, respectively). This may be indicative of changes in bacterial carbon/nutrient stoichiometry with changes in substrate stoichiometry¹², and therefore also of a more diffuse shift between the two system states than indicated by the dotted vertical line drawn in Fig. 3.

In the Arctic, concentrations of labile DOC in river input can be large¹³, reaching levels comparable in magnitude to the sum of DOC added throughout our experimental period in the high-level (3C and 3CSi) treatments. However, the ‘more organic carbon gives less organic carbon’ effect occurs at low glucose addition rates. Our results are therefore also relevant to environments without extreme inputs. Although our experiment mimics the situations with allochthonous input of organic carbon, we do, however, see its primary importance in the illustration of how internal trophic control mechanisms of the microbial food web work as a function of changing stoichiometric ratios in the dissolved pool.

The present indication is that natural systems may exist in both of the two states of bacterial growth rate limitation^{14–16}. Our experiment demonstrated large differences in both fate and effects of degradable organic carbon, depending on these states. The general principles of the trophic interactions involved should apply to both marine and limnic¹⁷ pelagic systems at any latitude, but one aspect is particularly relevant to the Arctic: any analysis of the relationship between

temperature and bacterial growth would be very different, depending on whether bacterial growth is controlled by the system’s production of labile DOC, or by the parameters determining bacterial success in algal–bacterial competition for nutrients.

With changes expected in ice cover and density profiles in the Arctic Ocean¹⁸, temperature, light conditions and vertical stability are likely to change, potentially leading to altered rate and composition of nutrient supply to the photic zone, and thus altered conditions for algal–bacterial competition. The differences between the +Si and –Si gradients in our experiment highlight the complex role of silicate in these interactions. With depth of the silica-cline sensitive to the sinking characteristics of the diatom community¹⁹, and a role of silicate in the balance between carbohydrate-producing *Phaeocystis* and diatoms²⁰, silicate would seem to have a large potential for changing the ratio between availability of organic carbon and nitrogen to the heterotrophic prokaryotic community. Models aspiring to capture present and to predict future carbon cycles in the Arctic therefore clearly require a correct understanding of the stoichiometric coupling between carbon and limiting nutrients, not only in the autotrophic but also in the heterotrophic parts of the food web.

METHODS SUMMARY

Commercial, 1 m³, translucent polyethylene tanks were used as mesocosm containers. For details of the mesocosm set-up, see Supplementary Information. Gross production, respiration and net community production were measured as oxygen changes in light–dark, 24-h-incubated, bottles. Chlorophyll was measured fluorometrically, TOC was measured by high-temperature catalytic oxidation, and leucine incorporation was measured by the centrifugation method as incorporation of ³H-leucine in 1.5-ml samples. Bacterial abundance was counted by flow cytometry.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 17 March; accepted 4 July 2008.

Published online 20 August 2008.

- Riebesell, U. *et al.* Enhanced biological carbon consumption in a high CO₂ ocean. *Nature* **450**, 545–548 (2007).
- delGiorgio, P. A., Cole, J. J. & Cimleris, A. Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* **385**, 148–151 (1997).
- Thingstad, T. F., Hagstrom, A. & Rassoulzadegan, F. Accumulation of degradable DOC in surface waters: is it caused by a malfunctioning microbial loop? *Limnol. Oceanogr.* **42**, 398–404 (1997).
- Thingstad, T. F., Nielsen, T. G., Hansen, A. S. & Levensen, H. Control of bacterial production in cold waters. A theoretical analysis of mechanisms relating bacterial production and zooplankton biomass in Disko Bay, Western Greenland. *Mar. Ecol. Prog. Ser.* **228**, 15–24 (2002).
- Bersheim, K. Y. *et al.* Photosynthetic algal production, accumulation and release of phytoplankton storage carbohydrates and bacterial production in a gradient in daily nutrient supply. *J. Plankton Res.* **27**, 743–755 (2005).
- Pomeroy, L. R. & Deibel, D. Temperature regulation of bacterial activity during the spring bloom in Newfoundland coastal waters. *Science* **233**, 359–361 (1986).
- Rose, J. M. & Caron, D. A. Does low temperature constrain the growth rates of heterotrophic protists? Evidence and implications for algal blooms in cold waters. *Limnol. Oceanogr.* **52**, 886–895 (2007).
- Stroeve, J., Holland, M. M., Meier, W., Scambos, T. & Serreze, M. Arctic sea ice decline: faster than forecast. *Geophys. Res. Lett.* **34**, L24501 (2007).
- Bellerby, R. G. J., Olsen, A., Furevik, T. & Anderson, L. A. in *Climate Variability in the Nordic Seas* (eds Drange, H., Dokken, T. M., Furevik, T., Gerdes, R. & Berger, W.) Geophysical Monograph Series 189–198 (AGU, 2005).
- Ducklow, H. W. *et al.* Constraining bacterial production, conversion efficiency and respiration in the Ross Sea, Antarctica, January–February, 1997. *Deep-Sea Res. Part II-Top. Stud. Oceanogr.* **47**, 3227–3247 (2000).
- Fagerbakke, K. M., Heldal, M. & Norland, S. Content of carbon, nitrogen, oxygen, sulfur and phosphorus in native aquatic and cultured bacteria. *Aquat. Microb. Ecol.* **10**, 15–27 (1996).
- Martinussen, I. & Thingstad, T. F. Utilization of N, P, and organic C by heterotrophic bacteria. II. Comparison of experiments and a mathematical model. *Mar. Ecol. Prog. Ser.* **37**, 285–293 (1987).
- Holmes, R. M. *et al.* Lability of DOC transported by Alaskan rivers to the arctic ocean. *Geophys. Res. Lett.* **35**, L03402 (2008).
- Graneli, W., Carlsson, P. & Bertilsson, S. Bacterial abundance, production and organic carbon limitation in the Southern Ocean (39–62 degrees S, 4–14 degrees E) during the austral summer 1997/1998. *Deep-Sea Res. Part II-Top. Stud. Oceanogr.* **51**, 2569–2582 (2004).

15. Church, M. J., Hutchins, D. A. & Ducklow, H. W. Limitation of bacterial growth by dissolved organic matter and iron in the Southern Ocean. *Appl. Environ. Microbiol.* **66**, 455–466 (2000).
16. Rivkin, R. & Anderson, M. Inorganic nutrient limitation of oceanic bacterioplankton. *Limnol. Oceanogr.* **42**, 730–740 (1997).
17. Stets, E. G. & Cotner, J. B. The influence of dissolved organic carbon on bacterial phosphorous uptake and bacteria-phytoplankton dynamics in two Minnesota lakes. *Limnol. Oceanogr.* **53**, 137–147 (2008).
18. Carmack, E. & Wassmann, P. Food webs and physical-biological coupling on pan-Arctic shelves: Unifying concepts and comprehensive perspectives. *Prog. Oceanogr.* **71**, 446–477 (2006).
19. Rey, F. & Skjoldal, H. R. Consumption of silicic-acid below the euphotic zone by sedimenting diatom blooms in the Barents Sea. *Mar. Ecol. Prog. Ser.* **36**, 307–312 (1987).
20. Wassmann, P. *et al.* Spring bloom development in the marginal ice zone and the central Barents Sea. *PSZN I: Mar. Ecol.* **20**, 321–346 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was financed by the Research Council of Norway through the International Polar Year project 175939/S30 'PAME-Nor' (IPY activity ID no. 71), with additional support from the strategic institution project 158936/I10 'Patterns in microbial diversity', Bjerknes Centre of Climate Research Centre of Excellence Project 146003/V30, project 178441/S40 'Interact' and project 184860/S30 'MERCLIM'. Support was received also from Norsk Hydro Produksjon AS project number 5404889, and from the Svalbard Science Forum as 'Arktisstipend'. We thank Kings Bay A/S and the staff at Ny Ålesund for help with logistics.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.F.T. (frede.thingstad@bio.uib.no).

METHODS

Nutrients. Soluble reactive phosphorous and silicate were determined according to refs 21 and 22, respectively. Ammonium was measured fluorometrically using a Turner Designs Model 10-AU equipped with optical kit number 10-303, following a previously reported protocol²³. Owing to analysis failure there are no ammonium data available from day 9 onwards. Nitrate was determined on stored samples fixed with chloroform (approximately 200 µl per 20 ml sample) and kept refrigerated until measured by an autoanalyser. No detectable nitrate was found (data not shown).

Turnover times. Uptake rate of orthophosphate was measured using ³³P-orthophosphate (ref. 24). ³³P-orthophosphate (Amersham, 370 MBq ml⁻¹) was added to samples at a final concentration of 125 pmol l⁻¹. Samples for the subtraction of the background and abiotic adsorption were fixed with 100% trichloroacetic acid (TCA) before isotope addition. Samples were incubated in the dark at *in situ* temperature. The incubation time varied between 20 s and 190 min (short enough to assure a linear relationship between the fraction of isotope adsorbed and the incubation time, but long enough to reliably detect isotope uptake above background levels). Incubation was stopped by a cold chase of 100 mmol l⁻¹ KH₂PO₄ (final concentration, 1 mmol l⁻¹). Samples were filtered onto 25-mm polycarbonate filters with 0.2-µm pore sizes, which were placed on a Millipore 12-place manifold with Whatman (GF/C) glass fibre filters saturated with 100 mmol l⁻¹ KH₂PO₄ as support. After filtration, filters were placed in polyethylene scintillation vials with Counting cocktail (Ecoscint A, National Diagnostics), and were radio-assayed. After the radioactivities of the filter were corrected for those of the blank filter obtained from fixed samples, $T_{[PO_4]}$ was calculated as $T_{[PO_4]} = -t/\ln(1 - f)$, where f is the fraction (no dimension) of added isotope collected on the 0.2 µm filter after the incubation time (t).

Uptake rate of glucose was measured using ¹⁴C-glucose (ref. 25, modified by ref. 26). D-[U-¹⁴C]-glucose (Amersham, 7.4 MBq ml⁻¹) was added to samples at a final concentration of 100 nmol l⁻¹. After 1 h of incubation in the dark at *in situ* temperature, the sample was split into two. Particulate ¹⁴C (>0.2 µm) uptake was measured on 10 ml samples filtered on 0.2-µm pore-size cellulose nitrate filters, and ¹⁴C-CO₂ was absorbed on 25 mm Whatman (GF/F) glass fibre filters with 250 µl phenethylamine fixed inside the cap of 20-ml polyethylene scintillation vials containing 10 ml. Filters were placed in polyethylene scintillation vials with Counting cocktail (Ecoscint A, National Diagnostics) and were radio-assayed. Turnover time of glucose was calculated as the inverse of the fraction of added isotope consumed per hour.

Chlorophyll *a* was measured fluorometrically²⁷. Alternating every second day, triplicate samples or a single sample of 100 ml water from each tank was filtered onto Nucleopore filters (47 mm diameter, 0.22 µm pore size). The filters were extracted in 90% acetone, at 4 °C for 10–12 h, before analysis on a Turner Designs 10-AU Fluorometer calibrated with pure chlorophyll *a* (Sigma Chemicals Inc.).

Net community production and community respiration were measured as oxygen production and consumption using light/dark 125-ml biological oxygen demand (BOD) bottles incubated *in situ* for 24 h. Gross primary production was calculated by subtraction. Two light and two dark bottles were used for each measurement. Oxygen concentration was measured before and after incubation in all bottles using the OxyMini optode system (World Precision Instruments) with the O₂-sensitive membrane glued to the inside wall of the BOD bottles. To get the integrated 24 h response to the nutrient additions, samples for these measurements were collected after addition of nutrients and mixing of the mesocosms. Assuming ammonium to be the major nitrogen source, oxygen rates were converted to carbon rates using a photosynthetic quotient of 1 (ref. 28).

Total bacterial and viral numbers were determined using the FACS Calibur flow cytometer (Becton Dickinson) equipped with an air-cooled laser providing 15 mW at 488 nm with standard filter set-up. The samples were fixed with glutaraldehyde at a final concentration of 0.5%. For bacterial counts the samples were stored at 4 °C until analysis the same day, whereas viral samples were fixed for 30 min at 4 °C, snap-frozen and stored in liquid nitrogen until further analysis. Each sample was diluted from 10- to 1,000-fold (two to three different dilutions) in 0.2 µm filtered 1× TE buffer and stained with SYBR Green I (Molecular Probes Inc.) for 10 min in the dark, with bacteria at room temperature and viruses at 80 °C (ref. 28). The samples were analysed for 1 min at a flow rate of around 50 µl min⁻¹ and events rates between 50 and 1,000 with the discriminator set on green fluorescence. Discrimination of bacteria and virus populations was based on groups observed in scatter plots of side scatter signal versus green DNA dye fluorescence (SYBR Green I).

Rates of bacterial production were determined by incorporation of tritium-labelled leucine³⁰ using the centrifugation procedure of ref. 31. Triplicate samples and one prefixed control sample were incubated with ³H-Leu (4,27 TBq mmol⁻¹, Perkin Elmer) at a final concentration of 60 nM. Incubation was performed in the dark at *in situ* temperature for 1 h. Samples were fixed with 5% TCA, final concentration, and then centrifuged at 14,000g for 10 min before removal of the supernatant. The samples were washed twice by adding 5% TCA, vortex, centrifugation and removal of supernatant. Counting cocktail (Ecoscint A, National Diagnostics) was added and the uptake of radioactive leucine was measured by scintillation counting.

Total organic carbon was measured using high temperature catalytic oxidation³².

21. Koroleff, F. in *Methods of Seawater Analysis* 2nd revised and extended edn (eds Grasshoff, K., Ehrhardt, M. & Kremling, K.) 125–131 (Chemie, 1983).
22. Valderrama, J. C. in *Manual on Harmful Marine Microalgae* IOC Manuals and Guides number 33 (eds Hallegraeff, G. M., Anderson, D. M. & Cembella, A. D.) 262–265 (UNESCO, 1995).
23. Holmes, R. M., Aminot, A., Kéroul, R., Hooker, A. H. & Peterson, B. J. A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can. J. Aquat. Sci.* **56**, 1801–1808 (1999).
24. Thingstad, T. F., Skjoldal, E. F. & Bohne, R. A. Phosphorus cycling and algal-bacterial competition in Sandsfjord, western Norway. *Mar. Ecol. Prog. Ser.* **99**, 239–259 (1993).
25. Hobbie, J. E. & Crawford, C. C. Bacterial uptake of organic substrate: new methods of study and application to eutrophication. *Verh. Internat. Verein. Limnol.* **17**, 725–730 (1969).
26. Havskum, H. et al. Silicate and labile DOC interfere in structuring the microbial food web via algal-bacterial competition for mineral nutrients: Results of a mesocosm experiment. *Limnol. Oceanogr.* **48**, 129–140 (2003).
27. Parsons, T. R., Maita, Y. & Lalli, C. M. *A Manual of Chemical and Biological Methods for Seawater Analysis* 107–110 (Pergamon, 1984).
28. Eriksen, N. T., Riisgard, F. K., Gunther, W. S. & Iversen, J. J. L. On-line estimation of O₂ production, CO₂ uptake, and growth kinetics of microalgal cultures in a gas-tight photobioreactor. *J. Appl. Phycol.* **19**, 161–174 (2007).
29. Marie, D., Brussaard, C. P. D., Thyraug, R., Bratbak, G. & Vaulot, D. Enumeration of marine viruses in culture and natural samples by flow cytometry. *Appl. Environ. Microbiol.* **65**, 45–52 (1999).
30. Simon, M. & Azam, F. Protein content and protein synthesis rates of planktonic marine bacteria. *Mar. Ecol. Prog. Ser.* **51**, 201–213 (1989).
31. Smith, D. & Azam, F. A simple, economical method for measuring bacterial protein synthesis in seawater using ³H-leucine. *Mar. Microb. Food Webs* **6**, 107–114 (1992).
32. Børsheim, K. Y. Bacterial production rates and concentrations of organic carbon at the end of the growing season in the Greenland Sea. *Aquat. Microb. Ecol.* **21**, 115–123 (2000).

Platelet-derived growth factor- α receptor activation is required for human cytomegalovirus infection

Liliana Soroceanu¹, Armin Akhavan¹ & Charles S. Cobbs^{1,2}

Human cytomegalovirus (HCMV) is a ubiquitous human herpesvirus that can cause life-threatening disease in the fetus and the immunocompromised host¹. Upon attachment to the cell, the virus induces robust inflammatory, interferon- and growth-factor-like signalling^{2–9}. The mechanisms facilitating viral entry and gene expression are not clearly understood⁴. Here we show that platelet-derived growth factor- α receptor (PDGFR- α) is specifically phosphorylated by both laboratory and clinical isolates of HCMV in various human cell types, resulting in activation of the phosphoinositide-3-kinase (PI(3)K) signalling pathway. Upon stimulation by HCMV, tyrosine-phosphorylated PDGFR- α associated with the p85 regulatory subunit of PI(3)K and induced protein kinase B (also known as Akt) phosphorylation, similar to the genuine ligand, PDGF-AA. Cells in which PDGFR- α was genetically deleted¹⁰ or functionally blocked were non-permissive to HCMV entry, viral gene expression or infectious virus production. Re-introducing human *PDGFRA* gene into knockout cells restored susceptibility to viral entry and essential viral gene expression. Blockade of receptor function with a humanized PDGFR- α blocking antibody (IMC-3G3)¹¹ or targeted inhibition of its kinase activity with a small molecule (Gleevec)¹² completely inhibited HCMV viral internalization and gene expression in human epithelial, endothelial and fibroblast cells. Viral entry in cells harbouring endogenous PDGFR- α was competitively inhibited by pretreatment with PDGF-AA. We further demonstrate that HCMV glycoprotein B directly interacts with PDGFR- α , resulting in receptor tyrosine phosphorylation, and that glycoprotein B neutralizing antibodies¹³ inhibit HCMV-induced PDGFR- α phosphorylation. Taken together, these data indicate that PDGFR- α is a critical receptor required for HCMV infection, and thus a target for novel anti-viral therapies.

HCMV, a β -herpesvirus, is the most common cause of congenital infection and an important pathogen in immunocompromised individuals¹. Viral attachment elicits a potent cellular interferon-like response^{2,5,6,9,14} which activates downstream growth-factor-like receptor tyrosine kinase (RTK) and integrin signalling pathways^{4,15}. HCMV modulation of the PI(3)K/Akt pathway is an important mechanism of apoptotic inhibition, ensuring long-term virus survival¹⁶. Prior evidence suggested that HCMV activation of human epidermal growth factor receptor (EGFR), in conjunction with integrin co-receptors, facilitates activation of downstream signalling molecules such as PI(3)K/Akt, phospholipase C γ and focal adhesion kinase^{15,17,18}. However, we and others demonstrated that EGFR is not required for cellular expression of HCMV-essential genes, or for virus-induced signalling^{19,20}. Instead, we found that upon short-term infection of human cells, HCMV caused phosphorylation of an approximately 180-kDa protein, distinct from EGFR, that could be co-immunoprecipitated with the p85 regulatory subunit of PI(3)K¹⁹.

Therefore, we hypothesized that another, as yet undiscovered, RTK might be activated by HCMV and mediate viral entry and expression.

To identify this putative RTK, we used a human phospho-specific RTK antibody array to screen human embryonic lung fibroblasts (HELs) that were either mock- or HCMV (Towne strain)-infected for 10 min. Only PDGFR- α was highly tyrosine phosphorylated upon infection with HCMV (Fig. 1a). Western blot analyses of these same protein lysates using a different phospho-specific antibody for PDGFR- α corroborated this observation (Fig. 1b). Independent quantitative enzyme-linked immunosorbent assays (ELISAs) confirmed PDGFR- α phosphorylation by Towne, AD169 and TR²¹ strains (Fig. 1c). HCMV did not induce tyrosine phosphorylation of the related RTK, PDGFR- β (Supplementary Fig. 1). Ultraviolet-inactivated HCMV induced PDGFR- α phosphorylation, whereas virus inactivated by heat did not (Supplementary Fig. 2).

To determine whether the HCMV-induced PDGFR- α phosphorylation was cell-type specific, we used U87 glioma (neuro-epithelial origin), HEL (fibroblast) and human umbilical vein endothelial cells (HUVECs, mesenchymal origin). In all three cell types, infection with HCMV Towne induced PDGFR- α phosphorylation, similar to the genuine ligand (Fig. 1d).

Based on these data, we hypothesized that PDGFR- α was the approximately 180 kDa protein we previously identified associated with the p85 regulatory subunit of PI(3)K, upon HCMV attachment¹⁹. To confirm this, we conducted co-immunoprecipitation experiments in HCMV or mock-infected cells. Immunoblotting of p85^{PI(3)K} immunoprecipitated proteins and whole-cell lysates with antibodies specific for PDGFR- α and phosphotyrosine indicated that PDGFR- α was tyrosine phosphorylated and associated with p85^{PI(3)K} upon HCMV short-term stimulation (Fig. 2a). Specificity of PDGFR- α phosphorylation by the Towne, AD169 and TR strains was examined in the presence or absence of IMC-3G3, a humanized PDGFR- α blocking antibody (ImClone¹¹). Pretreatment with IMC-3G3 significantly inhibited PDGFR- α phosphorylation induced by all HCMV strains tested (Fig. 2b). To investigate HCMV-induced activation of the downstream PI(3)K/Akt signalling pathway, we used both the IMC-3G3 antibody and a PDGFR- α kinase inhibitor, imatinib mesylate (Gleevec¹²), to block either PDGFR- α binding or its activation. Akt phosphorylation induced by either HCMV or PDGF-AA was abolished by IMC-3G3 and Gleevec (Fig. 2c). An isotype-matched negative control antibody did not inhibit HCMV- or PDGF-induced PDGFR- α or Akt phosphorylation (not shown). Thus, blocking PDGFR- α or inhibiting its activity prevents HCMV-mediated activation of the PI(3)K/Akt signalling pathway, an important pathway in the HCMV viral life cycle²².

To determine whether PDGFR- α is critical for viral internalization and gene expression, we used a well-characterized viral entry assay to measure internalization of the pp65 viral tegument protein after HCMV-treated cells were shifted from 4 °C to 37 °C²³. We used

¹Department of Neurosciences, California Pacific Medical Center Research Institute, Suite 220, 475 Brannan Street, San Francisco, California 94107, USA. ²Department of Neurological Surgery, University of California, San Francisco, 787 Moffitt, 505 Parnassus Avenue, San Francisco, California 94143, USA.

human and murine cells engineered to knockout, knockdown or overexpress PDGFR- α . As shown in Fig. 3a, 60 min after shifting to 37 °C, HEL cells demonstrate evidence of viral internalization, indicated by immunostaining of nuclear pp65 in over 65% of cells per microscopic field (top row). Short interfering RNA (siRNA)-mediated knockdown of PDGFR- α (Fig. 3a, lowest two rows) caused near-complete blockade of viral internalization, compared with the non-targeting, control siRNA-treated cells ($P < 0.001$, Fig. 3b). Similarly, murine fibroblasts obtained from embryos of *Pdgfra* knockout mice¹⁰ (embryonic lethal) showed no pp65-positive nuclei (Fig. 3a, second row) whereas in fibroblasts from parental strain (325S) over 25% of cells were pp65 positive (Fig. 3a, third row). Re-introducing human *PDGFRA* into the knockout cells restored and augmented HCMV internalization in these cells (Fig. 3a, fourth row, 80% of cells are pp65 positive) even compared with positive control cells (Fig. 3b, $P < 0.01$, Student's *t*-test).

To determine whether genetic ablation of PDGFR- α prevents cellular expression of essential HCMV gene products, we measured

expression of IE1 (*UL123*) in murine cells, as well as IE1 and pp65 (*UL83*) in human cells after infection with HCMV. Viral gene expression was undetectable in the PDGFR- α knockout murine cells (Fig. 3c) and in the human HEL cells pretreated with *PDGFRA* siRNA (Fig. 3d), compared with controls. Prolonged activation of human PDGFR- α by HCMV resulted in downregulation of receptor levels, consistent with a recent report²⁴ (Fig. 3d, upper panel, lane 2, in control siRNA-treated cells). IE1 expression was also undetectable after infection of the PDGFR- α null mouse fibroblasts with the AD169 strain, indicating that this effect is not strain specific (data not shown). Pretreatment with PDGFR- α blocking agents IMC-3G3 antibody and Gleevec completely inhibited HCMV IE1 protein expression in human HEL and U87 glioma cells (Fig. 3e) as well as in HUVECs (Supplementary Fig. 3).

We next investigated whether PDGFR- α expression was required for production of infectious virus using siRNA knockdown of PDGFR- α in HEL cells and a Towne-green fluorescent protein (GFP)-expressing virus²⁵ for visualization of infected cells and plaque formation. Forty-eight to sixty hours after siRNA transfection, HEL cells were exposed to HCMV for 1 h and monitored daily under a fluorescence microscope. Duplicate cultures were used for IE1 staining at 12 h after infection, whereas a third set of cultures was used to measure plaque formation (Supplementary Figs 4 and 5). Six days after infection, supernatants of these cells were used to infect naive HEL cells and assess production of infectious virus (Fig. 3f and Supplementary Fig. 4). We found near-complete inhibition of both

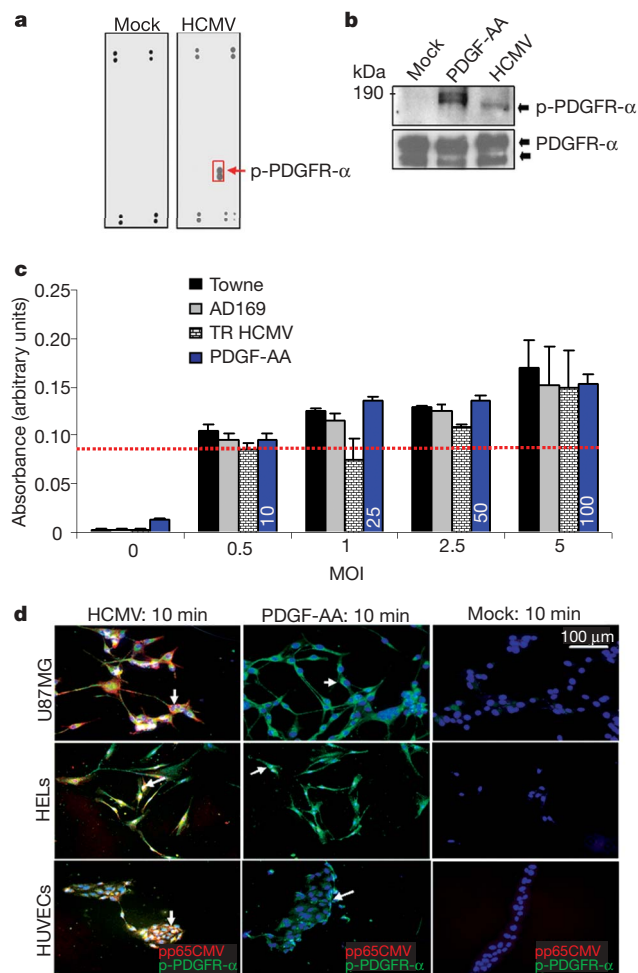


Figure 1 | HCMV induces tyrosine phosphorylation of human PDGFR- α .

a, Lysates of mock- or HCMV-treated cells were hybridized to a human phospho-RTK array. HCMV phosphorylates PDGFR- α (arrow). **b**, Western blot of HELs stimulated with mock, HCMV, or PDGF-AA, with indicated antibodies. **c**, Phospho-PDGFR- α -specific ELISA of HELs stimulated with indicated HCMV strains and PDGF-AA (dotted line corresponds to 4,000 pg ml⁻¹ phospho-PDGFR- α). Mean absorbance values ($n = 6$) \pm s.d. are shown. **d**, Immunofluorescence of pp65 and phospho-PDGFR- α in indicated cells stimulated with mock, HCMV or PDGF. Nuclei are stained with 4,6-diamidino-2-phenylindole (DAPI).

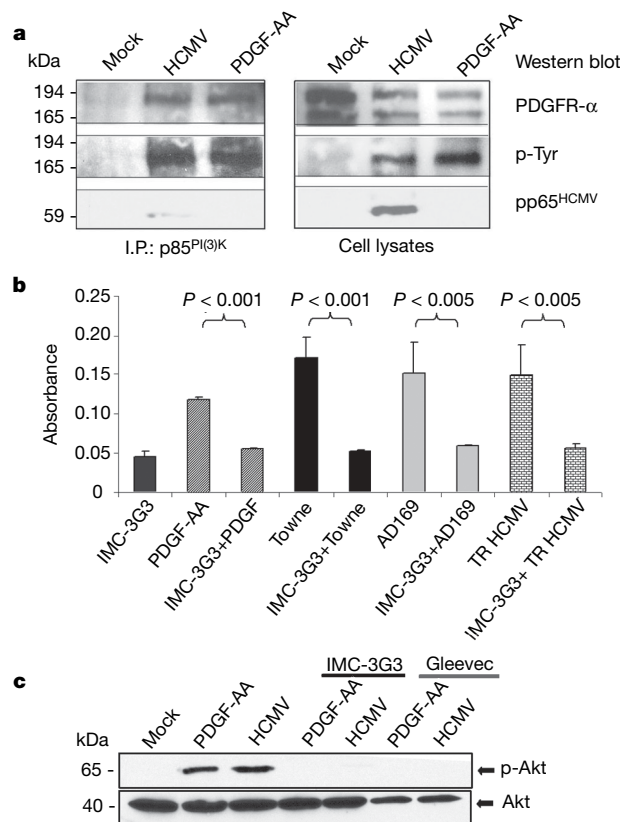


Figure 2 | HCMV activates the PI(3)K/Akt pathway in a PDGFR- α -dependent manner.

a, HEL cells stimulated with HCMV or PDGF-AA were subjected to immunoprecipitation and western blot analyses with indicated antibodies. **b**, Phospho-PDGFR- α ELISA of cells with or without pretreatment with IMC-3G3 (10 μ g ml⁻¹, 2 h) followed by HCMV (MOI = 1) or PDGF (10 ng ml⁻¹) for 10 min. Mean values ($n = 6$) \pm s.d. are shown. **c**, Western blot of HEL mock, HCMV or PDGF-AA stimulated (10 min), with or without IMC-3G3 (2 μ g ml⁻¹) or Gleevec (100 nM). The same membrane was used for p-Akt^{Ser 473} and Akt.

viral gene expression and infectious virus production in cells that did not express PDGFR- α at the time of infection. Plaque formation in HEL cells was also completely blocked by PDGFR- α knockdown (Supplementary Fig. 5).

To determine the relative importance of PDGFR- α versus downstream PI(3)K activation for viral gene expression and infectious virus production, we performed a series of experiments using *p110 α ^{PI(3)K}* (also known as *PIK3CA*) siRNA and non-targeting siRNA-treated cells. Suppression of the PI(3)K pathway by p110 α knockdown resulted in a delay in viral gene expression, yet allowed viral entry and infectious virus production, albeit at lower levels than controls, which is in agreement with previous studies using PI(3)K inhibitors²² (Supplementary Fig. 6). These data indicate that although PI(3)K activation is important for the HCMV life cycle, expression of a functional PDGFR- α is essential.

We next tested whether the authentic PDGFR- α ligand inhibits viral entry. Pretreatment with PDGF-AA significantly decreased HCMV entry in HEL cells, suggesting that PDGF-AA competes with an HCMV protein (Supplementary Fig. 7). Because HCMV envelope

glycoprotein B (*UL55*) mediates viral entry and cellular signalling^{26,27}, we investigated whether glycoprotein B is the viral moiety directly interacting with PDGFR- α . Using a modified attachment assay, we found that a purified glycoprotein B peptide was internalized in mouse cells overexpressing human PDGFR- α , but not in PDGFR- α null cells (Fig. 4a). To demonstrate a direct interaction between PDGFR- α and glycoprotein B, we performed co-immunoprecipitation experiments using HEL cells that endogenously express PDGFR- α and purified recombinant full-length glycoprotein B²⁸. Reciprocal immunoblot analyses (Fig. 4b, c) demonstrate that glycoprotein B and PDGFR- α co-immunoprecipitate, indicating direct association between PDGFR- α and glycoprotein B as a bona fide mechanism for attachment/internalization of HCMV into the host cells. Using phosphor-PDGFR- α ELISA and western blot approaches, we found that full-length glycoprotein B induced PDGFR- α phosphorylation (Fig. 4d, e). Furthermore, two different glycoprotein B neutralizing antibodies¹³ significantly inhibited HCMV-induced PDGFR- α tyrosine phosphorylation (Fig. 4d, e), indicating that the PDGFR- α -glycoprotein B interaction is functionally relevant. Isotype control

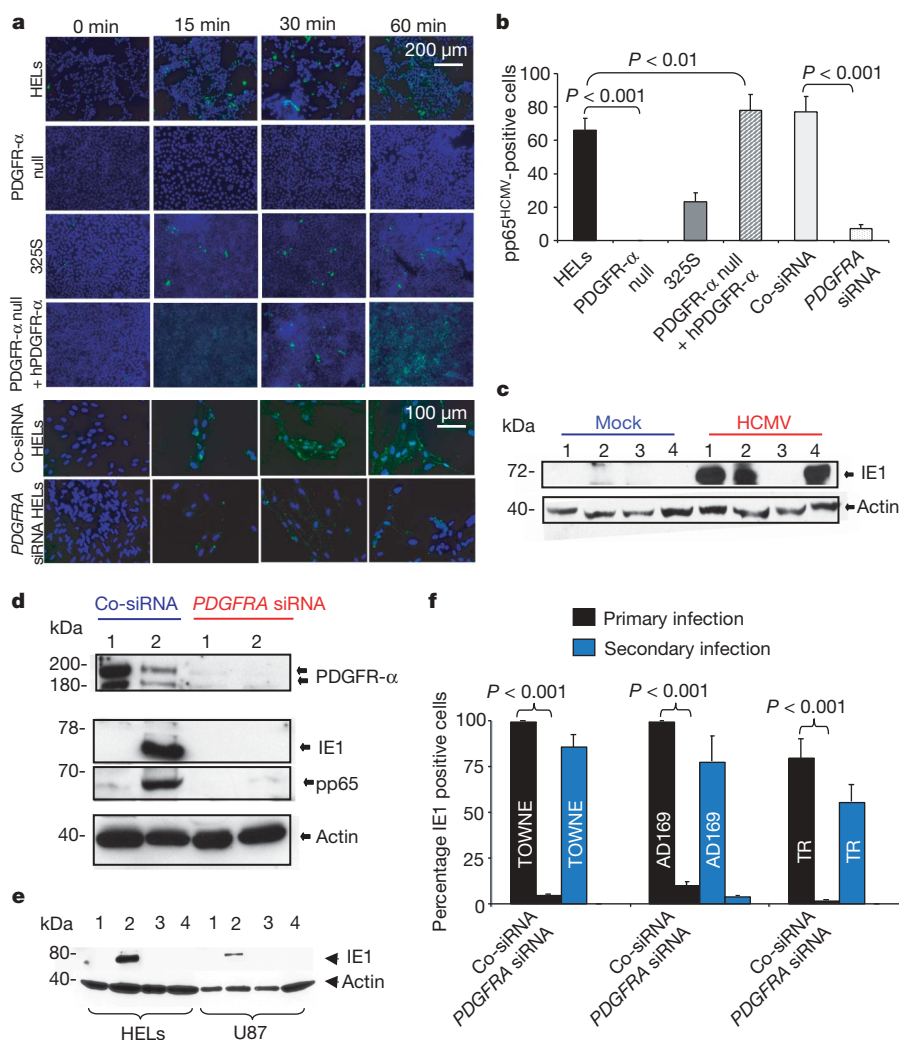


Figure 3 | Human PDGFR- α is required for HCMV entry, IE1 expression and infectious virus production. **a**, pp65 immunofluorescence after HCMV treatment (1 h, 4 °C) followed by shifting to 37 °C for indicated times. Rows represent (top to bottom): HELs, PDGFR- α null fibroblasts, parental fibroblasts, null fibroblasts overexpressing hPDGFR- α , HELs transfected with control or *PDGFR- α* siRNA. Nuclei were stained with DAPI. **b**, Average \pm s.d. pp65 positive per 100 cells counted in triplicate from **a**. **c**, Mock- or HCMV-infected cells were analysed by western blotting with the indicated antibodies. Lanes 1–4 indicate HELs, 325S, PDGFR- α null and PDGFR- α null overexpressing hPDGFR- α , respectively. **d**, HELs transfected

with control or *PDGFR- α* siRNA were mock (lanes 1) or HCMV- treated (lanes 2) and subjected to western blots with indicated antibodies. **e**, Western blot of HEL and U87 lysates after infection with mock (lanes 1), HCMV (lanes 2) or HCMV pretreated with IMC-3G3 (10 μ g ml⁻¹, 12 h, lanes 3) or Gleevec (100 nM, 1 h, lanes 4). **f**, *PDGFR- α* siRNA-treated cells infected with indicated HCMV strains and immunostained for IE1 12 h after infection (primary infection). Six days after infection, supernatants were used to infect naive HEL cells, followed by IE1 immunostaining and quantification (secondary infection). Average ($n = 6$) values \pm s.d. are shown.

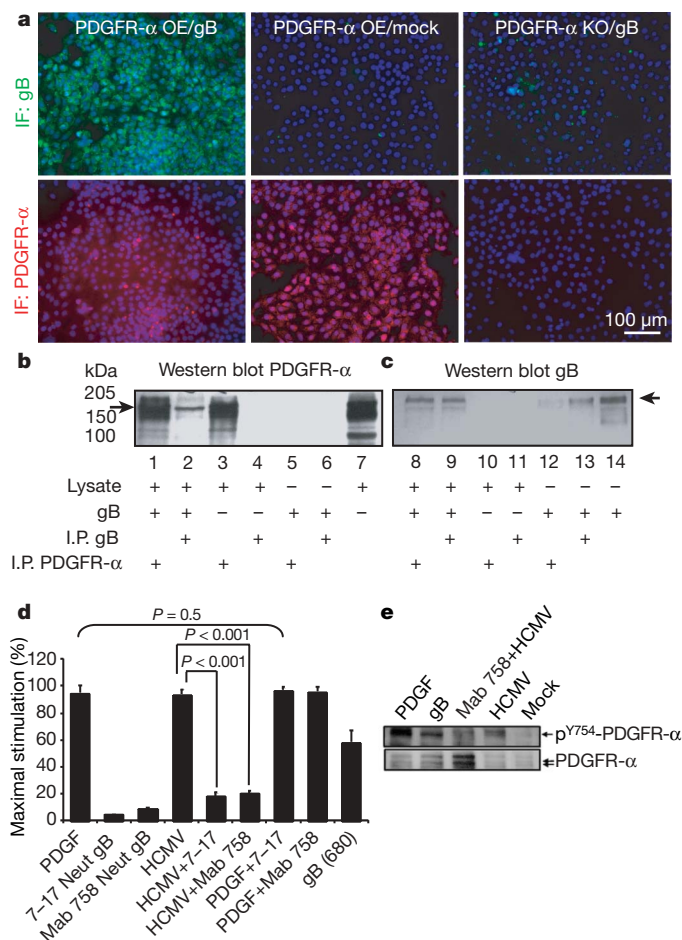


Figure 4 | HCMV Glycoprotein B binds and activates PDGFR- α .

a, Immunofluorescence detection of glycoprotein B (upper panels) and PDGFR- α (lower panels) in PDGFR- α knockout (KO) cells and PDGFR- α overexpressing (OE) cells, after incubation with the glycoprotein B peptide or mock treatment. Reduced PDGFR- α surface staining in glycoprotein-B-treated cells is likely due to receptor internalization.

b, c, Immunoprecipitation of PDGFR- α and glycoprotein B from HEL lysates and full-length soluble glycoprotein B alone or pre-incubated together. Immunoprecipitates were subjected to western blot with the indicated antibodies. **d**, ELISA measurements of human PDGFR- α phosphorylation after stimulation with HCMV (MOI = 1), PDGF (10 ng ml⁻¹) or recombinant soluble glycoprotein B (30 μ g ml⁻¹) in the presence or absence of glycoprotein B neutralizing antibodies 7-17 and Mab 758 (5 μ g ml⁻¹); bars, \pm s.d. **e**, Portions of the same lysates used in **d** were analysed by western blot with indicated antibodies.

antibodies or neutralizing antibodies against other viral glycoproteins (glycoproteins H and N) did not prevent HCMV-induced receptor activation (not shown).

Overall, data presented here indicate HCMV requires PDGFR- α binding and activation for viral internalization, expression of essential viral genes, production of infectious virus and activation of downstream PI(3)K/Akt signalling. These findings do not exclude a potentially important role for other co-receptors during HCMV internalization and expression, such as integrin receptors. We demonstrate that viral interaction with PDGFR- α is facilitated by direct binding of the viral glycoprotein B to PDGFR- α . We further show that blockade of the PDGFR- α receptor pathway with pharmaceutical agents currently in human use may prove a powerful antiviral strategy for the management of HCMV-related disease. Because both PDGFR- α and HCMV play important roles in the pathophysiology of human development, inflammation, vascular disease and cell-proliferative disorders, an increased understanding of their interaction may elucidate novel molecular mechanisms underlying these conditions.

METHODS SUMMARY

Cells and viruses. PDGFR- α knockout and parental 325S mouse fibroblasts were a gift from M. Tallquist (Southwestern University¹⁰). Although HCMV is a human-specific virus and cannot be propagated in murine cells, HCMV internalization and HCMV immediate early viral gene expression occur in murine cells²⁹. cDNA for human PDGFRA was obtained from C. Heldin (Uppsala University). HCMV strains AD169 and Towne (American Type Culture Collection, ATCC) were propagated in HELs for less than five passages. HELs were infected at a multiplicity of infection (MOI) of 1 (in serum-free media) and cell supernatant was collected over a period of 5–7 days after infection, when the cytopathic effect was about 100%. Virus-containing media were first centrifuged (1,500g) to remove cell debris and further concentrated using a sucrose gradient centrifugation (80,000g, 4 °C) as described³⁰. Virus stock aliquots were kept at –80 °C. Mock controls were generated in parallel by conditioning and processing uninfected cultures identically. TR HCMV clinical isolate and GFP-CMV were obtained from W. Britt.

siRNA experiments. HELs were transfected with 100 nM of either the ‘smart Pool’ siRNA to human PDGFRA, human p110 $\alpha^{PI(3)K}$ or the non-targeting siRNA pool (Dharmacon) using standard Lipofectamine 2000 reagent protocol (Invitrogen). Seventy-two hours after transfection, siRNA-transfected cells were HCMV- or mock-infected. PDGFR- α or p110 $\alpha^{PI(3)K}$ expression levels were measured using standard western blot and immunofluorescence analyses (antibodies from Cell Signalling).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 May; accepted 25 June 2008.

Published online 13 August 2008.

- Britt, W. J. & Alford, C. A. *Fields Virology* 3rd edn (Raven Press, 1996).
- Andreoni, K. A., Wang, X., Huang, S. M. & Huang, E. S. Human cytomegalovirus hyperimmune globulin not only neutralizes HCMV infectivity, but also inhibits HCMV-induced intracellular NF- κ B, Sp1, and PI3-K signalling pathways. *J. Med. Virol.* **67**, 33–40 (2002).
- Boyle, K. A., Pietropaolo, R. L. & Compton, T. Engagement of the cellular receptor for glycoprotein B of human cytomegalovirus activates the interferon-responsive pathway. *Mol. Cell. Biol.* **19**, 3607–3613 (1999).
- Compton, T. Receptors and immune sensors: the complex entry path of human cytomegalovirus. *Trends Cell Biol.* **14**, 5–8 (2004).
- Netterwald, J. R. et al. Postattachment events associated with viral entry are necessary for induction of interferon-stimulated genes by human cytomegalovirus. *J. Virol.* **78**, 6688–6691 (2004).
- Ozato, K., Tailor, P. & Kubota, T. The interferon regulatory factor family in host defense: mechanism of action. *J. Biol. Chem.* **282**, 20065–20069 (2007).
- Simmen, K. A. et al. Global modulation of cellular transcription by human cytomegalovirus is initiated by viral glycoprotein B. *Proc. Natl Acad. Sci. USA* **98**, 7140–7145 (2001).
- Yurochko, A. D. et al. Induction of the transcription factor Sp1 during human cytomegalovirus infection mediates upregulation of the p65 and p105/p50 NF- κ B promoters. *J. Virol.* **71**, 4638–4648 (1997).
- Zhu, H., Cong, J. P. & Shen, T. Use of differential display analysis to assess the effect of human cytomegalovirus infection on the accumulation of cellular RNAs: induction of interferon-responsive RNAs. *Proc. Natl Acad. Sci. USA* **94**, 13985–13990 (1997).
- Andrews, A. et al. Platelet-derived growth factor plays a key role in proliferative vitreoretinopathy. *Invest. Ophthalmol. Vis. Sci.* **40**, 2683–2689 (1999).
- Loizos, N. et al. Targeting the platelet-derived growth factor receptor alpha with a neutralizing human monoclonal antibody inhibits the growth of tumor xenografts: implications as a potential therapeutic target. *Mol. Cancer Ther.* **4**, 369–379 (2005).
- Sandler, C. et al. Imatinib mesylate inhibits platelet derived growth factor stimulated proliferation of rheumatoid synovial fibroblasts. *Biochem. Biophys. Res. Commun.* **347**, 31–35 (2006).
- Britt, W. J. Neutralizing antibodies detect a disulfide-linked glycoprotein complex within the envelope of human cytomegalovirus. *Virology* **135**, 369–378 (1984).
- Waltenberger, J. et al. Different signal transduction properties of KDR and Flt1, two receptors for vascular endothelial growth factor. *J. Biol. Chem.* **269**, 26988–26995 (1994).
- Feire, A. L., Koss, H. & Compton, T. Cellular integrins function as entry receptors for human cytomegalovirus via a highly conserved disintegrin-like domain. *Proc. Natl Acad. Sci. USA* **101**, 15470–15475 (2004).
- Cooray, S. The pivotal role of phosphatidylinositol 3-kinase-Akt signal transduction in virus survival. *J. Gen. Virol.* **85**, 1065–1076 (2004).
- Wang, X., Huang, D. Y., Huang, S. M. & Huang, E. S. Integrin $\alpha_3\beta_3$ is a coreceptor for human cytomegalovirus. *Nature Med.* **11**, 515–521 (2005).
- Wang, X. et al. Epidermal growth factor receptor is a cellular receptor for human cytomegalovirus. *Nature* **424**, 456–461 (2003).
- Cobbs, C. S. et al. Human cytomegalovirus induces cellular tyrosine kinase signaling and promotes glioma cell invasiveness. *J. Neurooncol.* **85**, 271–280 (2007).

20. Isaacson, M. K., Feire, A. L. & Compton, T. Epidermal growth factor receptor is not required for human cytomegalovirus entry or signaling. *J. Virol.* **81**, 6241–6247 (2007).
21. Murphy, E. *et al.* Coding potential of laboratory and clinical strains of human cytomegalovirus. *Proc. Natl Acad. Sci. USA* **100**, 14976–14981 (2003).
22. Johnson, R. A. *et al.* Human cytomegalovirus up-regulates the phosphatidylinositol 3-kinase (PI3-K) pathway: inhibition of PI3-K activity inhibits viral replication and virus-induced signaling. *J. Virol.* **75**, 6022–6032 (2001).
23. English, E. P., Chumanov, R. S., Gellman, S. H. & Compton, T. Rational development of beta-peptide inhibitors of human cytomegalovirus entry. *J. Biol. Chem.* **281**, 2661–2667 (2006).
24. Gredmark, S. *et al.* Human cytomegalovirus downregulates expression of receptors for platelet-derived growth factor by smooth muscle cells. *J. Virol.* **81**, 5112–5120 (2007).
25. Murphy, E. A., Streblow, D. N., Nelson, J. A. & Stinski, M. F. The human cytomegalovirus IE86 protein can block cell cycle progression after inducing transition into the S phase of permissive cells. *J. Virol.* **74**, 7108–7118 (2000).
26. Boyle, K. A. & Compton, T. Receptor-binding properties of a soluble form of human cytomegalovirus glycoprotein B. *J. Virol.* **72**, 1826–1833 (1998).
27. Carlson, C., Britt, W. J. & Compton, T. Expression, purification, and characterization of a soluble form of human cytomegalovirus glycoprotein B. *Virology* **239**, 198–205 (1997).
28. Wang, Z. *et al.* Recombinant modified vaccinia virus Ankara expressing a soluble form of glycoprotein B causes durable immunity and neutralizing antibodies against multiple strains of human cytomegalovirus. *J. Virol.* **78**, 3965–3976 (2004).
29. Thomas, B. *Viruses and the Cellular Immune Response* (Marcel Dekker, 1993).
30. Huang, E. S., Chen, S. T. & Pagano, J. S. Human cytomegalovirus. I. Purification and characterization of viral DNA. *J. Virol.* **12**, 1473–1481 (1973).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Tallquist for the PDGFR- α knockout mouse fibroblasts, C. Heldin for the human PDGFR- α cDNA, D. Diamond for providing the soluble glycoprotein B, and N. Loizos (ImClone) for the IMC-3G3 antibody. We are grateful to W. Britt (University of Alabama at Birmingham) for viruses, glycoprotein B neutralizing antibodies and discussions. This study was supported by an institutional grant from California Pacific Medical Center Research Institute and by the Arthur Flaming Foundation.

Author Contributions L.S. and A.A. performed experiments; L.S., A.A. and C.S.C. designed experiments, analysed data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.C. (charles.cobbs@gmail.com).

METHODS

Cell culture, plasmids, transfection and additional HCMV strains. HELs, U87 glioma and HUVECs were obtained from ATCC and maintained in DMEM plus 10% FCS, except for HUVECs which were grown in endothelial cell media (Cascade Biologicals), plus growth factors. Mouse PDGFR- α knockout cells were transfected with human *PDGFRA* cDNA using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Towne-GFP (from W. Britt, University of Alabama at Birmingham) is a recombinant HCMV strain that expresses GFP under the early promoter *UL127*, as previously described³¹. Virus titres were determined by IE1 immunohistochemical staining, as previously described³². Optimization of siRNA cell delivery was performed by co-transfecting with the targeting siRNA pools (for example, *PDGFRA* siRNA) and fluorescently labelled oligonucleotides (siGLO-RISC free, Dharmacon) that localize to the nucleus and allow assessment of uptake into cells. Targeting siRNA oligonucleotides and siGLO were mixed 1:1 (50 nM each) with different amounts of Lipofectamine 2000. Twenty-four hours later, cells were fixed, counterstained with DAPI and counted. The average number of green fluorescent cells (a measure of transfection efficiency) was between 72% and 84% (from a total of 100% DAPI-positive nuclei).

siRNA sequences. The sequences are listed 5'–3'. PI(3)K *p110 α* : GCGAAUUCACACUAUU; GUGGUAAGUCCAGAUUA; GCUUAGA GUUGGAGUUUGA; GACCCUAGCCUUAAGAUAAA. Human *PDGFRA*: CG AGACUCCUGUAACCUUAUU; GAGCUUCACCUAUCAGUUUU; GACAG UGGCAUUAACUAUU; GAAUAGGGUAGCUUCCUGUU. Non-targeting sequences: UGGUUUACUAGUCGACUAA; UGGUUUACAUGUUUUCUGA; UGGUUUACAUGUUGUGUGA; UGGUUUACAUGUUUUCUAA.

Western blot, immunoprecipitation and immunofluorescence analyses. For stimulation experiments, cells were serum-starved for 24 h, followed by stimulation with HCMV (MOI = 0.5), PDGF-AA (5–10 ng ml⁻¹, R&D Systems), or mock, for 10 min. In some cases, cells were pretreated with IMC-3G3 (N. Loizos, ImClone) at 10 μ g ml⁻¹ (2 or 12 h) or Gleevec (100 nM, 1 h) before HCMV exposure. Cell lysates were analysed by SDS–polyacrylamide gel electrophoresis (SDS–PAGE). Antibodies used were as follows: polyclonal anti-PDGFR- α (1/500), phosphor-tyrosine clone 4G10 (1/1,000) and monoclonal anti-p85 PI(3)K (1/1,000) (Upstate Biotechnology), anti HCMV IE1 MAB810 (Chemicon; 1/1,000), anti HCMV pp65 (Novocastra, 1/1,000), anti-phosphor-PDGFR- α (1/500, pTyr 754), anti-phosphor-Akt (Ser 473, 1/1,000) and total Akt (1/1,000; all from Cell Signalling). Anti-actin polyclonal control antibody was used (1/500, Sigma). Immunoprecipitations were performed using protein G (Pierce) according to the manufacturer's instructions. For immunofluorescence, we used monoclonal HCMV pp65 and phosphor-PDGFR- α (p-Tyr 754, Santa Cruz Biotechnology), overnight at 4 °C, followed by incubation for 1 h with secondary antibodies conjugated to Alexa 488, or Alexa 568 (1/5,000, Molecular Probes). Nuclei were counterstained with DAPI. Co-immunoprecipitation experiments were performed using full-length soluble purified recombinant glycoprotein B (glycoprotein B 680, from D. Diamond, City of Hope, California) and detergent-soluble extracts of HEL cells generated in lysis buffer (1% NP-40, 75 mM NaCl and 50 mM Tris-HCl). Protein (500 μ g) from total cell extract was incubated with 25 μ g of glycoprotein B 680 in lysis buffer in the presence of an anti-glycoprotein B (Virusys) or anti-PDGFR- α (R&D) antibody overnight at 4 °C. Immune complexes were recovered with protein A–Sepharose beads (2 h incubation), denatured and separated on SDS–PAGE for western blot with anti-glycoprotein B (Virusys) or anti-PDGFR- α (Cell Signaling Technology) antibodies.

Viral attachment and internalization assays. Mouse fibroblasts and HEL cells (72 h after siRNA transfections) were incubated with HCMV (MOI = 0.5) or mock treated for 1 h at 4 °C, after which cells were returned to 37 °C for 15, 30 or 60 min. Cells were fixed using methanol (20 min) and processed for pp65 immunofluorescence. Four low-power fields were counted for each condition and pp65 immunoreactive cells were recorded for each 100 cells. Internalization assays were repeated twice.

HCMV glycoprotein B (peptide) attachment assays. Glycoprotein B peptide binding experiments were performed similar to viral attachment assays. After incubation for 1 h with the glycoprotein B peptide (100 nM, 4 °C), cells were

returned to 37 °C for 60 min, washed, fixed and processed for double immunofluorescence for glycoprotein B (1 μ g ml⁻¹, monoclonal antibody, Virusys) and PDGFR- α (2 μ g ml⁻¹, Upstate). The glycoprotein B peptide (Ray Biotech) contains amino acids 27–84 from the AD169 strain and 27–81 from the Towne strain.

Human phospho-PDGFR- α ELISA. ELISA for human phospho-PDGFR- α was performed with a kit (R&D, catalogue number DYC2114-2). HEL cells were grown in 24-well plates (40,000 cells per ml) and serum-starved 48 h before short-term (10 min) stimulation with various agents, as described in Figs 2b and 4d. Lysis of cells was done as per kit instructions. The capture antibody was a mouse anti-human PDGFR- α ; anti-phosphotyrosine-HRP antibody was used for detection. Recombinant human phosphorylated PDGFR- α was used as a positive control. Reaction products were read using a microplate reader set at 450 nm. All samples were run in triplicate and each experiment was repeated at least twice. For receptor-blocking experiments, cells were pretreated with IMC-3G3 (10 μ g ml⁻¹, 12 h) or Gleevec (100 nM, 1 h). To test the effects of glycoprotein B neutralizing antibodies, HCMV (MOI = 1) was pre-incubated in serum-free media with antibodies specific for glycoprotein B 7–17, MAB 758 (ref. 33) or control isotype-matched antibodies (5 μ g ml⁻¹) for 1 h before cell stimulation. Additional antibodies tested included neutralizing antibody against HCMV glycoproteins N (ref. 34) and H (ref. 35).

Measurements of infectious virus production. HEL cells (where indicated treated with siRNA) were infected with HCMV Towne or CMV-GFP for 1 h, washed and grown for 6 days at 37 °C. A duplicate set of cultures was analysed by immunofluorescence at 12 h after infection to assess the percentage of IE1 positive cells, as a measure of 'primary infection'. Six days after infection, supernatant from these cells were centrifuged to exclude cell debris and used to infect naive HEL cultures as described above. Cells at 12 h after infection were stained for IE1, and IE1-positive cells were counted among a total of 100 cells per low-magnification microscopic field, four fields per condition. These counts were used to determine the level of 'secondary infection': that is, infectious virus production from the primary infected cells. Where indicated, CMV-GFP infected cells were monitored daily under a fluorescence microscope. Each condition was assayed in triplicate.

Plaque formation assays. Plaque formation was assayed as previously described³⁶. Briefly, confluent HEL cells in six-well cluster plates were incubated with CMV-GFP (MOI = 1, 1 h) in 0.5 ml growth media. Cells were washed and returned to 37 °C (in complete growth media). Twenty-four hours later, supernatant was harvested and used to infect naive HEL cultures, which were monitored for plaque formation for 6–14 days. Plaque formation was photographed daily using an inverted fluorescence microscope. At day 14, plaques in ten low-magnification fields/conditions were counted (each experimental condition was tested in six independent wells).

Statistical analyses. A two-tailed paired Student's *t*-test was used to compare data sets and obtain *P* values for all comparisons; *P* values are indicated on figure panels.

- Isomura, H. & Stinski, M. F. The human cytomegalovirus major immediate-early enhancer determines the efficiency of immediate-early gene transcription and viral replication in permissive cells at low multiplicity of infection. *J. Virol.* **77**, 3602–3614 (2003).
- Chan, G., Stinski, M. F. & Guilbert, L. J. Human cytomegalovirus-induced upregulation of intercellular cell adhesion molecule-1 on villous syncytiotrophoblasts. *Biol. Reprod.* **71**, 797–803 (2004).
- Britt, W. J., Jarvis, M. A., Drummond, D. D. & Mach, M. Antigenic domain 1 is required for oligomerization of human cytomegalovirus glycoprotein B. *J. Virol.* **79**, 4066–4079 (2005).
- Shimamura, M., Mach, M. & Britt, W. J. Human cytomegalovirus infection elicits a glycoprotein M (gM)/gN-specific virus-neutralizing antibody response. *J. Virol.* **80**, 4591–4600 (2006).
- Li, L., Coelingh, K. L. & Britt, W. J. Human cytomegalovirus neutralizing antibody-resistant phenotype is associated with reduced expression of glycoprotein H. *J. Virol.* **69**, 6047–6053 (1995).
- Mar, E. C., Cheng, Y. C. & Huang, E. S. Effect of 9-(1,3-dihydroxy-2-propoxymethyl)guanine on human cytomegalovirus replication *in vitro*. *Antimicrob. Agents Chemother.* **24**, 518–521 (1983).

LETTERS

Autophagy in thymic epithelium shapes the T-cell repertoire and is essential for tolerance

Jelena Nedjic^{1†}, Martin Aichinger^{1†}, Jan Emmerich^{1†}, Noboru Mizushima² & Ludger Klein^{1,3}

Recognition of self-antigen-derived epitopes presented by major histocompatibility complex class II (MHC II) molecules on thymic epithelial cells (TECs) is critical for the generation of a functional and self-tolerant CD4 T-cell repertoire. Whereas haematopoietic antigen-presenting cells generate MHC-II–peptide complexes predominantly through the processing of endocytosed polypeptides¹, it remains unknown if and how TECs use unconventional pathways of antigen presentation. Here we address the role of macroautophagy, a process that has recently been shown to allow for endogenous MHC II loading^{2–6}, in T-cell repertoire selection in the mouse thymus. In contrast to most other tissues, TECs had a high constitutive level of autophagy. Genetic interference with autophagy specifically in TECs led to altered selection of certain MHC-II-restricted T-cell specificities and resulted in severe colitis and multi-organ inflammation. Our findings indicate that autophagy focuses the MHC-II–peptide repertoire of TECs on their intracellular milieu, which notably comprises a wide array of otherwise strictly ‘tissue-specific’ self antigens^{7,8}. In doing so, it contributes to T-cell selection and is essential for the generation of a self-tolerant T-cell repertoire.

Thymic epithelial cells (TECs) are the only non-haematopoietic cell types that constitutively express MHC II. Cortical TECs (cTECs) have predominantly been implicated in positive selection of thymocytes carrying self-MHC-restricted T-cell receptors (TCRs)⁹, whereas medullary TECs (mTECs) fulfil a non-redundant role in tolerance induction by ‘promiscuously’ expressing otherwise tissue-restricted antigens (TRAs)^{7,8}. Notably, tolerance induction towards mTEC-derived antigens can operate through direct antigen recognition on mTECs¹⁰ or by cross-presentation by thymic dendritic cells¹¹.

MHC-II–peptide ligands are primarily generated through processing of exogenous proteins¹. TECs, however, show a remarkably poor efficacy in capturing and processing extracellular antigens¹², suggesting alternative pathways of MHC II loading. Candidate routes through which endogenous loading of MHC II molecules may be achieved are a TAP-dependent pathway¹³, chaperone-mediated autophagy¹⁴ and macroautophagy^{2–6}.

Macroautophagy is a bulk protein degradation process that is upregulated to sustain metabolic fitness during food deprivation¹⁵. Autophagosomes fuse with lysosomes for degradation of their cargo, and it is conceivable that at this point an intersection with the MHC II pathway occurs². Because the thymus has been identified as a site of unusual starvation-independent autophagic activity¹⁶, we hypothesized that this phenomenon may facilitate endogenous MHC II loading in TECs and thus contribute to positive selection and/or tolerance induction.

Using green fluorescent protein (GFP)–LC3 autophagy reporter mice¹⁶ (Supplementary Fig. 1a), we confirmed that the thymus had

robust, starvation-independent autophagic activity (Fig. 1a). In contrast, in many other tissues—for example, muscle—detectable autophagy is induced only upon starvation (Fig. 1a and Supplementary Fig. 1b). Absence of punctae in thymi of reporter mice with a targeted disruption of *Atg5* (autophagy-related gene 5)^{17,18}, an essential component of autophagosome formation, indicated that these structures were not unspecific aggregates (Supplementary Fig. 1c).

When purified thymic stromal cells were analysed (Fig. 1b), the highest frequency ($68.5 \pm 2.9\%$) of autophagy-positive cells (≥ 5

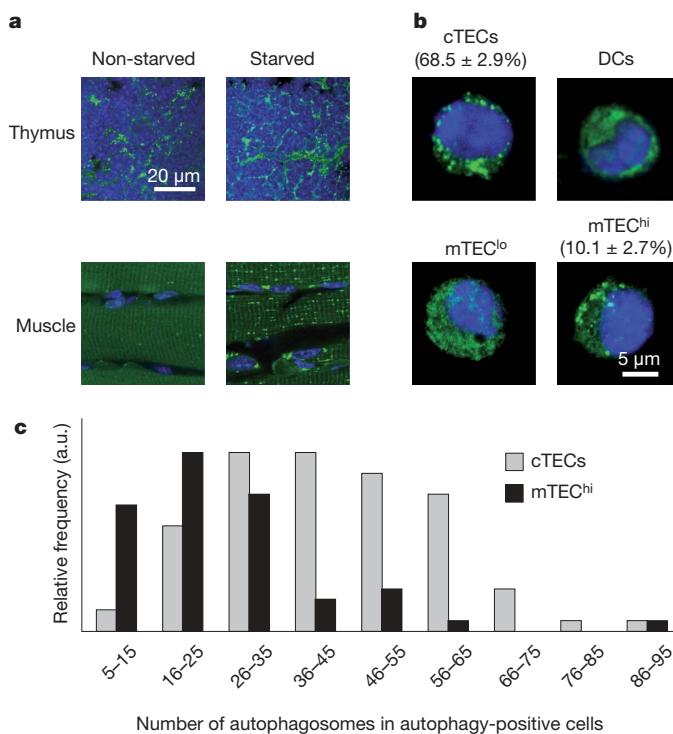


Figure 1 | Constitutive autophagy in TECs. **a**, Sections of the thymus and muscle of fed and starved GFP–LC3 mice were stained with anti-GFP Alexa 488 antibody (green) and 4,6-diamidino-2-phenylindole (DAPI; blue). Images are representative of $n = 6$. **b**, Thymic stromal cells from GFP–LC3 mice were fixed onto glass slides and stained with anti-GFP Alexa 488 and DAPI. Typical autophagy-positive cTECs and mTEC^{hi} are shown. The average frequencies \pm s.d. of cTECs and mTEC^{hi} scoring autophagy positive (that is, ≥ 5 autophagosomes) in two independent preparations are indicated. Less than 1% of dendritic cells (DCs) and mTEC^{lo} were positive. **c**, Relative distribution of the number of autophagosomes in autophagy-positive cTECs and mTEC^{hi}; a.u., arbitrary units.

¹Research Institute of Molecular Pathology, Doktor Bohr Gasse 7, 1030 Vienna, Austria. ²Department of Physiology and Cell Biology, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8613, Japan. ³Institute for Immunology, Ludwig-Maximilians-University, Goethestrasse 31, 80336 Munich, Germany. [†]Present addresses: Institute for Immunology, Ludwig-Maximilians-University, Goethestrasse 31, 80336 Munich, Germany (J.N., M.A.); Schering-Plough BioPharma (formerly DNAX, Inc.), 901 California Avenue, Palo Alto, California 94304, USA (J.E.).

autophagosomes) was found among cTECs. In mTEC subsets^{19,20}, $10.1 \pm 2.7\%$ of mature MHC II^{hi} CD80^{hi} mTECs (mTEC^{hi}) scored positive, whereas autophagy was negligible among immature MHC II^{lo} CD80^{lo} mTECs (mTEC^{lo}), demonstrating a co-segregation of autophagy with 'promiscuous' antigen expression²¹. The number of autophagosomes in positive cells was scattered around maxima of about 35 or 20 in cTECs or the mTEC^{hi} subset, respectively (Fig. 1c).

Although autophagy has been implicated in developmental processes^{15,22}, *Atg5*^{-/-} mice are born at Mendelian ratios without evident developmental abnormalities¹⁷. However, neonatal lethality of *Atg5*^{-/-} mice, at least in part caused by a perinatal metabolic crisis, precluded a direct assessment of T-cell selection and tolerance in this system. To assess the capacity of *Atg5*^{-/-} thymi to support T-cell development beyond the perinatal phase, we transplanted embryonic *Atg5*^{-/-} thymi under the renal capsule of normal adult recipients. When analysed after 6 to 8 weeks, *Atg5*^{-/-} thymi were significantly smaller than wild-type controls (Fig. 2a).

Several observations suggested that the reduced cellularity of *Atg5*^{-/-} grafts was not caused by aberrant differentiation of the epithelial compartment. Thus, *Atg5*^{-/-} grafts were properly compartmentalized into cortical and medullary areas (Fig. 2b) and contained cTECs and mTECs at normal frequencies (Supplementary Fig. 2). Expression of MHC I and II on *Atg5*^{-/-} epithelial cells was essentially normal; however, a slight decrease in MHC II density was observed on cTECs (Fig. 2c). *Atg5*^{-/-} mTECs showed the typical biphasic MHC II expression indicative of normal segregation into mTEC^{lo} and mTEC^{hi} subsets, and promiscuous antigen expression was intact (Fig. 2d).

Next we considered whether the reduced cellularity of *Atg5*^{-/-} grafts may have resulted from perturbed T-cell development. Ratios of thymocyte subsets including CD25⁺ Foxp3⁺ regulatory T

cells were normal in H-2^d *Atg5*^{-/-} (Fig. 3a) or H-2^b *Atg5*^{-/-} thymi (Supplementary Fig. 3a). TCR- α or TCR- β V-region use by CD4 single-positive (SP) thymocytes was also normal (Supplementary Fig. 3b, c), indicating that *Atg5*^{-/-} deficiency in TECs was compatible with development of a polyclonal T-cell repertoire containing the main T-cell lineages.

Positive selection is remarkably flexible regarding the requirements for the composition of MHC-peptide (MHCp) ligands on cTECs, as exemplified by the efficient selection of polyclonal CD4 T cells in mice expressing a markedly reduced spectrum of MHCp complexes^{23,24}. If autophagy were to contribute to the generation of MHC II ligands on TECs, its absence would be expected to have rather subtle consequences. Two predictions followed from these considerations. First, selection of particular monoclonal MHC-II-restricted TCR specificities should be affected in *Atg5*^{-/-} thymi. Second, if the reduced cellularity of *Atg5*^{-/-} thymi was caused by altered T-cell selection, a 'permissive' monoclonal TCR specificity should rescue the organ size. We tested these predictions by analysing the selection of several MHC-II-restricted transgenic TCRs (that is, TCR-HA restricted to I-E^d, AND restricted to I-E^k, DO11.10 restricted to I-A^d, and DEP and SEP restricted to I-A^b). In *Atg5*^{-/-} H-2^d lobes grafted into TCR-HA transgenic mice, both the frequency of CD4 SP cells expressing the TCR-HA receptor and the cellularity of lobes were significantly reduced as compared to controls (Fig. 3b). Similarly, *Atg5*^{-/-} H-2^b lobes grafted into SEP TCR transgenic hosts had reduced cellularity, and the decreased frequency of CD4 SP cells together with an increased expression of endogenous TCR- α chains indicated that there was selective pressure against this TCR (Supplementary Fig. 4). In contrast, the AND TCR seemed to be selected as efficiently in the *Atg5*^{-/-} to AND grafts as in the wild-type to AND controls; this finding coincided with equal cellularities (Fig. 3c). For the *Atg5*^{-/-} to DEP (Supplementary Fig. 5) and the *Atg5*^{-/-} to DO11.10 (Supplementary Fig. 6) transplants, we likewise observed efficient positive selection of the respective TCR and rescue of thymic cellularities. Of note, the frequency of clonotype-positive CD4 SP cells was significantly increased in the *Atg5*^{-/-} to DO11.10 lobes, indicating that *Atg5* deficiency could also be beneficial for the selection of particular specificities. Transplantation of *Atg5*^{-/-} lobes into mixed (DO11.10 *Rag2*^{-/-}/TCR-HA *Rag2*^{-/-}) bone marrow chimaeras accentuated the opposing consequences of epithelial *Atg5* deficiency for these TCRs (Supplementary Fig. 6d). In contrast to these observations with MHC-II-restricted TCRs, selection of three MHC-I-restricted transgenic TCRs (P14 restricted to H-2D^b, OT-I restricted to H-2K^b and HY restricted to H-2D^b) in *Atg5*^{-/-} or wild-type lobes was indistinguishable, both with respect to expression of the transgenic TCR chains and thymic cellularity (Fig. 3d and Supplementary Fig. 7; see also Supplementary Table 1 for a summary). Taken together, these findings were consistent with autophagy in TECs specifically shaping the MHC-II-restricted T-cell repertoire, probably through the generation of particular MHC-II-peptide ligands. This underscores the notion that distinct pathways of antigen presentation in TECs may be essential for the generation of the 'normal' T-cell repertoire^{25,26}.

Evidence that autophagy can indeed mould the composition of MHC II ligands was recently obtained by mass spectrometric analysis of MHC-II-bound peptides of B-lymphoblastoid cell lines in the presence or absence of autophagy⁶. Quantitative considerations rendered such a global assessment of the 'MHC ligandome' impossible for limited numbers of *ex vivo* isolated TECs. To visualize whether a particular MHC-class-II-bound peptide on cTECs was affected by genetic interference with autophagy, we used a monoclonal antibody (Y-Ae) that recognizes I-A^b when occupied by a I-E α -derived peptide²⁷. The I-E α ₅₂₋₆₈-I-A^b complex is abundant on haematopoietic antigen-presenting cells (~10% of all I-A^b-peptide complexes)²⁷, probably as a result of co-sorting of I-E and I-A molecules. Notably, however, this complex is comparatively underrepresented on cTECs, and it was hypothesized that this was due to competition

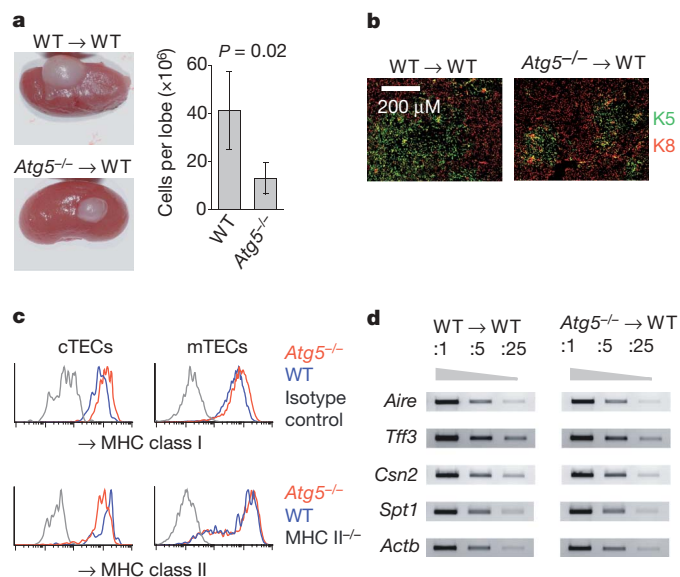


Figure 2 | Epithelial differentiation in the absence of *Atg5*. **a**, Thymi from wild-type (WT) or *Atg5*^{-/-} embryos were transplanted under the kidney capsule of wild-type recipients and analysed 6 weeks later. The average cellularities \pm s.d. are shown; $n = 5$. **b**, Normal compartmentalization of *Atg5*^{-/-} grafts in cytokeratin-8-positive (red; K8) cortical and cytokeratin-5-positive (green; K5) medullary regions; $n = 8$. **c**, Expression of MHC class I and II on cTECs or mTECs from *Atg5*^{-/-} to wild-type (WT) grafts (red), or wild-type to wild-type grafts (blue). Staining controls are depicted in grey; data are representative of three experiments. **d**, RT-PCR analysis of autoimmune regulator (*Aire*) and TRA expression in mTECs isolated from *Atg5*^{-/-} to wild-type, or wild-type to wild-type grafts. Expression of β -actin (*Actb*), intestinal trefoil factor (*Tff3*), casein 2 (*Csn2*) and salivary protein 1 (*Spt1*) are also shown (5-fold serial dilutions). Data are representative of two experiments.

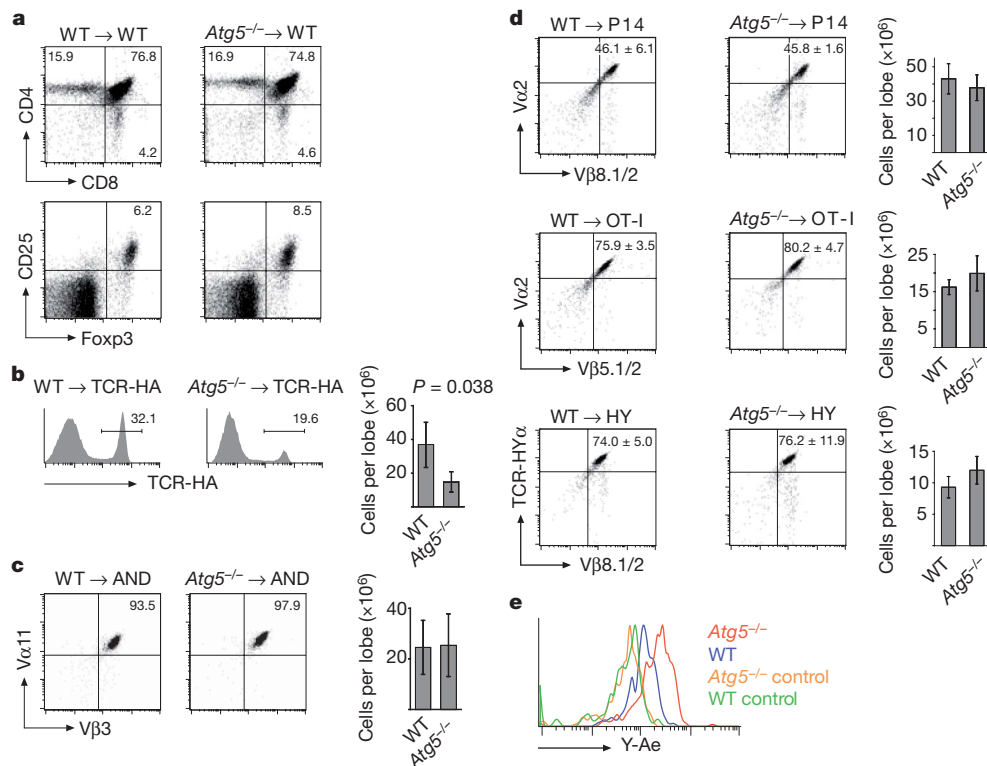


Figure 3 | *Atg5* deficiency modulates selection of MHC-II-restricted TCR specificities and the abundance of a specific MHCp complex. **a**, Polyclonal T-cell development in *Atg5*^{-/-} to wild-type (WT), or wild-type to wild-type grafts. The top dot plots show CD4 and CD8 staining; the bottom dot plots show expression of CD25 and Foxp3 by CD4 SP thymocytes; *n* ≥ 5. **b**, *Atg5*^{-/-} or wild-type lobes were grafted into TCR-HA transgenic recipients. Graft cellularity ± s.d. and the expression of the transgenic TCR on CD4 SP cells were analysed after 6 weeks. The frequency ± s.d. of TCR-HA⁺ cells among CD4 SP cells was 32.1% ± 3.2% in wild-type to TCR-HA, and 19.6% ± 4.3% in *Atg5*^{-/-} to TCR-HA lobes; *P* = 1.8 × 10⁻⁸ (*n* ≥ 14). **c**, *Atg5*^{-/-} or wild-type lobes were grafted into AND transgenic mice. Graft cellularity ± s.d. and the expression of the transgenic TCR (Vβ3-Vα11) on CD4 SP cells were analysed after 6 weeks. The frequency ± s.d. of Vβ3⁺ Vα11⁺ CD4 SP cells

was 93.5% ± 0.2% in wild-type to AND lobes (*n* = 3) and 97.9% ± 1.1% (*n* = 6) in *Atg5*^{-/-} to AND lobes (*P* = 9 × 10⁻⁴). **d**, Thymi from wild-type or *Atg5*^{-/-} embryos were grafted into P14, OT-I or HY TCR-transgenic recipients. The dot plots show the expression of the transgenic TCR V-segments on CD8 SP cells. Thymus cellularities ± s.d. are shown in the respective bar diagrams. Data are representative of *n* = 4 (P14), *n* ≥ 4 (OT-I) and *n* ≥ 2 (HY). **e**, Thymi from wild-type or (BALB/c × C57BL/6)_{F1} *Atg5*^{-/-} embryos were transplanted into (BALB/c × C57BL/6)_{F1} recipients. Six weeks after grafting, expression of the I-Eα₅₂₋₆₈-I-A^b complex on cTECs was analysed using the monoclonal antibody Y-Ae. *Atg5*^{-/-} cTECs are shown in red; wild-type cTECs are shown in blue. Control stainings are shown in orange (*Atg5*^{-/-}) and green (wild type). Data are representative of two independent experiments.

for I-A^b binding by peptides that may be expressed and/or processed in a cTEC-specific manner. When we analysed cTECs from F₁ (BALB/c × C57BL/6) *Atg5*^{-/-} lobes, we found an increased abundance of the I-Eα₅₂₋₆₈-I-A^b complex as compared to wild-type controls, whereas total MHC II levels showed an inverse behaviour (Fig. 3e, compare with Fig. 2c). These findings provided direct evidence that *Atg5* deficiency causes a quantitative shift in the MHC II ligandome of cTECs. We propose that in *Atg5*^{-/-} cTECs, a set of epitopes that rely on autophagy for endogenous MHC II loading may not compete with the Eα peptide, which itself may easily gain access to MHC II by other pathways.

To address whether the T-cell repertoire selected by *Atg5*^{-/-} TECs was self-tolerant, *Atg5*^{-/-} or *Atg5*^{+/-} thymi were grafted under the kidney capsule of athymic mice (designated *nu/nu*^{*Atg5*^{-/-}} or *nu/nu*^{WT}, respectively). The peripheral repertoire of CD4 T cells was similar with respect to Vα and Vβ use in both sets of chimaeras (Supplementary Fig. 8). However, the frequency of CD4 T cells with an activated phenotype (CD62L^{lo} and CD69⁺) was significantly increased in the *nu/nu*^{*Atg5*^{-/-}} chimaeras (Fig. 4a). Between 4 and 6 weeks after grafting, wasting began in *nu/nu*^{*Atg5*^{-/-}} but not *nu/nu*^{WT} chimaeras, and most of the *nu/nu*^{*Atg5*^{-/-}} chimaeras had to be killed between 8 to 16 weeks after grafting (Fig. 4b). Recipients of *Atg5*^{-/-} thymi had patches of flaky skin, and inspection of internal organs showed a massively enlarged colon, atrophy of the uterus, complete absence of fat pads and in many cases enlarged lymph nodes (Fig. 4c). Histological examination demonstrated inflammatory

infiltrates in the colon, liver, lung, uterus and Harderian gland (Fig. 4d and Supplementary Table 2). Chimaeras generated with 5 days deoxyguanosine treatment of embryonic lobes or by transplantation immediately after preparation yielded essentially identical outcomes, excluding organ-culture-related epithelial tissue damage or carry-over of *Atg5*^{-/-} haematopoietic cells as critical factors for disease development. The latter point was bolstered by the fact that *Atg5*^{-/-} to wild-type fetal liver chimaeras did not show any signs of autoimmunity (Supplementary Fig. 9; ref. 28).

Adoptive transfer of 1 × 10⁷ purified CD90⁺ T cells from *nu/nu*^{*Atg5*^{-/-}} chimaeras into *nu/nu* recipients recapitulated essentially all aspects of autoimmunity as observed in donor chimaeras (Supplementary Fig. 10). When transferred separately, CD4⁺ T cells from *nu/nu*^{*Atg5*^{-/-}} chimaeras elicited autoimmune symptoms more efficiently than CD8⁺ T cells, in line with *Atg5*^{-/-} deficiency in thymic epithelium primarily perturbing the selection of the CD4 T-cell repertoire (data not shown).

Taken together, we propose that perturbations in both positive and negative selection caused by *Atg5* deficiency in TECs may synergize to fuel autoimmunity. First, alterations in the composition of MHCp ligands on cTECs may affect positive selection, potentially creating a mildly lymphopenic environment that might foster autoimmunity. Second, impaired presentation of TRAs by mTECs may allow for exit into the periphery of autoreactive thymocytes that would normally be deleted or deviated into the regulatory T cell lineage.

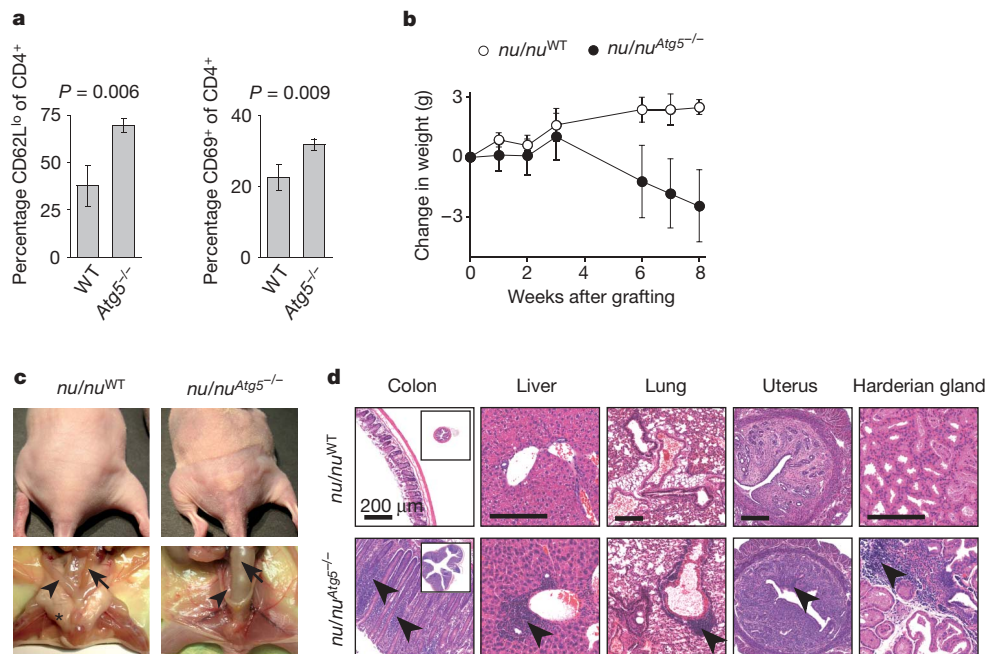


Figure 4 | *Atg5* deficiency in thymic epithelium causes colitis and multiorgan lymphoid infiltration. *Atg5*^{-/-} or wild-type (WT) lobes were grafted into athymic (*nu/nu*) mice (*nu/nu*^{*Atg5*^{-/-}} or *nu/nu*^{WT}, respectively). Data are representative of three experiments with $n \geq 4$. **a**, Expression of activation markers on lymph node CD4⁺ T cells in *nu/nu*^{*Atg5*^{-/-}} or *nu/nu*^{WT}

chimaeras. **b**, The changes in body weight over time are shown. **c**, The appearance of *nu/nu*^{*Atg5*^{-/-}} or *nu/nu*^{WT} chimaeras. The colon (arrow) and uterus (arrowhead) are highlighted; note the absence of fat pads (asterisk) in *nu/nu*^{*Atg5*^{-/-}} chimaeras. **d**, Haematoxylin and eosin staining of organs from *nu/nu*^{*Atg5*^{-/-}} or *nu/nu*^{WT} chimaeras. Arrows indicate infiltrates.

The organ selectivity of the autoimmune manifestations observed here remains to be explained. Of note, the Harderian gland itself has a substantial level of constitutive autophagy (Supplementary Fig. 11). Thus, autophagy-dependent epitope display in the periphery, if not counterbalanced by an analogous tolerogenic mechanism in the thymus, may be critical for immune-mediated tissue destruction. Along these lines, the human *ATG16L1* gene was recently identified as a susceptibility locus for Crohn's disease^{29,30}. Our data offer an explanation for the association of an autophagy-related gene with breakdown of intestinal immune homeostasis.

METHODS SUMMARY

Animals. All transgenic mouse strains have been described elsewhere. Animal protocols were approved by local authorities.

Antibodies and flow cytometry. Flow cytometry was performed according to standard procedures. See full Methods for antibodies.

GFP-LC3 analyses. Before analyses, mice were anaesthetized and perfused through the left ventricle with 4% paraformaldehyde. Organs were collected and fixed with sucrose. Tissue samples were embedded in optimal cutting temperature (OCT) medium and sectioned at 5 μ m before staining with an anti-GFP antibody. For quantification, purified thymic stromal cells were obtained from enzymatically digested thymi by a combination of density fractionation and FACS sorting. Stromal cells were fixed onto slides and processed for staining with an anti-GFP antibody.

Thymus transplantation. Embryonic thymi were transplanted under the kidney capsule of female animals. *Atg5*^{-/-} embryos and *Atg5*^{+/+} (wild-type) controls within individual experiments were obtained from the same *Atg5*^{+/+} pregnant female.

Histopathology. Organs were immersion-fixed and embedded in paraffin blocks. Sections were stained with haematoxylin and eosin and automatically scanned using the Zeiss MIRAX SCAN system.

Statistical analyses. All statistical analyses were performed using the two-tailed Student's *t*-test with unequal variance.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 19 May; accepted 26 June 2008.

Published online 13 August 2008.

1. Trombetta, E. S. & Mellman, I. Cell biology of antigen processing *in vitro* and *in vivo*. *Annu. Rev. Immunol.* **23**, 975–1028 (2005).

2. Schmid, D., Pypaert, M. & Munz, C. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* **26**, 79–92 (2007).
3. Paludan, C. *et al.* Endogenous MHC class II processing of a viral nuclear antigen after autophagy. *Science* **307**, 593–596 (2005).
4. Nimmerjahn, F. *et al.* Major histocompatibility complex class II-restricted presentation of a cytosolic antigen by autophagy. *Eur. J. Immunol.* **33**, 1250–1259 (2003).
5. Dorfel, D. *et al.* Processing and presentation of HLA class I and II epitopes by dendritic cells after transfection with *in vitro*-transcribed MUC1 RNA. *Blood* **105**, 3199–3205 (2005).
6. Dengjel, J. *et al.* Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. *Proc. Natl Acad. Sci. USA* **102**, 7922–7927 (2005).
7. Kyewski, B. & Klein, L. A central role for central tolerance. *Annu. Rev. Immunol.* **24**, 571–606 (2006).
8. Anderson, M. S. *et al.* Projection of an immunological self shadow within the thymus by the Aire protein. *Science* **298**, 1395–1401 (2002).
9. Starr, T. K., Jameson, S. C. & Hogquist, K. A. Positive and negative selection of T cells. *Annu. Rev. Immunol.* **21**, 139–176 (2003).
10. Aschenbrenner, K. *et al.* Selection of Foxp3⁺ regulatory T cells specific for self antigen expressed and presented by Aire⁺ medullary thymic epithelial cells. *Nature Immunol.* **8**, 351–358 (2007).
11. Gallegos, A. M. & Bevan, M. J. Central tolerance to tissue-specific antigens mediated by direct and indirect antigen presentation. *J. Exp. Med.* **200**, 1039–1049 (2004).
12. Klein, L., Roettinger, B. & Kyewski, B. Sampling of complementing self-antigen pools by thymic stromal cells maximizes the scope of central T cell tolerance. *Eur. J. Immunol.* **31**, 2476–2486 (2001).
13. Tewari, M. K., Sinnathamby, G., Rajagopal, D. & Eisenlohr, L. C. A cytosolic pathway for MHC class II-restricted antigen processing that is proteasome and TAP dependent. *Nature Immunol.* **6**, 287–294 (2005).
14. Zhou, D. *et al.* Lamp-2a facilitates MHC class II presentation of cytoplasmic antigens. *Immunity* **22**, 571–581 (2005).
15. Levine, B. & Klionsky, D. J. Development by self-digestion: molecular mechanisms and biological functions of autophagy. *Dev. Cell* **6**, 463–477 (2004).
16. Mizushima, N., Yamamoto, A., Matsui, M., Yoshimori, T. & Ohsumi, Y. *In vivo* analysis of autophagy in response to nutrient starvation using transgenic mice expressing a fluorescent autophagosome marker. *Mol. Biol. Cell* **15**, 1101–1111 (2004).
17. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432**, 1032–1036 (2004).
18. Mizushima, N. *et al.* Dissection of autophagosome formation using Atg5-deficient mouse embryonic stem cells. *J. Cell Biol.* **152**, 657–668 (2001).
19. Rossi, S. W. *et al.* RANK signals from CD4⁺3⁺ inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. *J. Exp. Med.* **204**, 1267–1272 (2007).

20. Gray, D., Abramson, J., Benoist, C. & Mathis, D. Proliferative arrest and rapid turnover of thymic epithelial cells expressing Aire. *J. Exp. Med.* **204**, 2521–2528 (2007).
 21. Derbinski, J. *et al.* Promiscuous gene expression in thymic epithelial cells is regulated at multiple levels. *J. Exp. Med.* **202**, 33–45 (2005).
 22. Huang, J. & Klionsky, D. J. Autophagy and human disease. *Cell Cycle* **6**, 1837–1849 (2007).
 23. Tourne, S. *et al.* Selection of a broad repertoire of CD4⁺ T cells in H-2Ma^{0/0} mice. *Immunity* **7**, 187–195 (1997).
 24. Ignatowicz, L., Kappler, J. & Marrack, P. The repertoire of T cells shaped by a single MHC/peptide ligand. *Cell* **84**, 521–529 (1996).
 25. Nakagawa, T. *et al.* Cathepsin L: critical role in li degradation and CD4 T cell selection in the thymus. *Science* **280**, 450–453 (1998).
 26. Murata, S. *et al.* Regulation of CD8⁺ T cell development by thymus-specific proteasomes. *Science* **316**, 1349–1353 (2007).
 27. Murphy, D. B. *et al.* A novel MHC class II epitope expressed in thymic medulla but not cortex. *Nature* **338**, 765–768 (1989).
 28. Pua, H. H., Dzhagalov, I., Chuck, M., Mizushima, N. & He, Y. W. A critical role for the autophagy gene Atg5 in T cell survival and proliferation. *J. Exp. Med.* **204**, 25–31 (2007).
 29. Hampe, J. *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in *ATG16L1*. *Nature Genet.* **39**, 207–211 (2007).
 30. Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* **39**, 596–604 (2007).
- Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.
- Acknowledgements** We acknowledge support from the Austrian National Science Fund (Sonderforschungsbereich F023 and grant Z58-B01 'Wittgenstein Prize Meinrad Busslinger'; to J.N. and L.K.) and the European Union (FP6 Integrated Project 'Eurothymaide'; Contract LSHB-CT-2003-503410; to M.A., J.E. and L.K.). Research at the Research Institute of Molecular Pathology is funded by Boehringer Ingelheim. We thank C. Spona for technical assistance and M. Busslinger, M. S. Anderson, B. Kyewski and J. Derbinski for comments on the manuscript. We also thank H. Ploegh for discussions in the earliest phase of this project.
- Author Contributions** J.N. was involved in all experiments (assisted by J.E., M.A. and L.K.). M.A. carried out the analysis of TRA expression. J.N., N.M. and L.K. designed experimental strategies. J.N. and L.K. wrote the manuscript. All authors discussed and commented on the contents of this manuscript.
- Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to L.K. (ludger.klein@med.lmu.de).

METHODS

Animals. *Atg5*^{-/-} and GFP-LC3 mice have been described previously^{17,16}. BALB/*c*^{nu/nu} animals were purchased from Taconic Farms. The AND TCR transgenic and B10.A animals were obtained from Jackson Laboratories. The TCR-HA, DO11.10, SEP and DEP TCR transgenic animals were described elsewhere^{31,32}. All mice were bred and maintained in individually ventilated cages in the animal facility of the Research Institute of Molecular Pathology under specific pathogen-free conditions. All animal studies were approved by local authorities (MA58) and were performed according to Austrian regulations.

Antibodies and flow cytometry. Biotin-conjugated monoclonal antibody Y-Ae was a gift from B. Kyewski. The following monoclonal antibodies were used: biotin-conjugated antibody to CD11c (HL3), CD62L (MEL14), CD80 (16-10A1), Vβ3 (KJ25) and Vβ14 (14-2); fluorescein isothiocyanate (FITC)-conjugated antibody to CD62L (MEL14), CD90.1 (HIS51), H-2K^b (AF6-88.5), Vα8.3 (KT50), Vβ3 (KJ25), Vβ4 (KT4), Vβ5.1.5.2 (MR9-4), Vβ6 (RR4-7), Vβ8.1/2 (MR5-2), Vβ8.3 (1B3.3) and Vβ13 (MR12-3); phycoerythrin-conjugated antibody to CD69 (H1.2F3), Ly51 (BP-1), Vα2 (B20.1), Vα11.1,11.2 (RR8-1), Vβ2 (B20.6), Vβ7 (TR310), Vβ8.3 (1B3.3) and Vβ10 (B21.5); CyChrome-conjugated antibody to CD8 (53-6.7) and CD45 (30-F11); phycoerythrin-Cy7 conjugated antibody to CD25 (PC61); allophycocyanin-conjugated antibody to CD4 (RM4-5) and CD8 (52-6.7); allophycocyanin-indotricarbocyanin-conjugated antibody to CD4 (GK1.5); and phycoerythrin-Cy7-conjugated streptavidin and FITC-conjugated streptavidin (Becton Dickinson). Phycoerythrin-conjugated monoclonal antibody to TCR-HYα (T3.70) was purchased from eBiosciences. *Ulex europaeus* agglutinin (UEA-1) was purchased from Sigma Aldrich and conjugated to Alexa 647 in our laboratory. Monoclonal antibodies specific for TCR-HA (6.5), DO11.10 (KJ1-26), EpCAM (G8.8), Vβ8 (F23.1), pan-MHC class II (P7.7) and CD4 (GK1.5) were purified from hybridoma supernatants and conjugated to biotin, Alexa Fluor 488 or Alexa Fluor 647 in our laboratory. Surface staining was performed according to standard procedures at a density of 1×10^6 to 3×10^6 cells per 50 μl. FoxP3 intracellular staining was carried out according to the manufacturer's recommendations with phycoerythrin- or allophycocyanin-conjugated monoclonal antibodies to mouse FoxP3 (FJK-16s, eBiosciences). A FACSCanto (Becton Dickinson) with FACSDiva software (Becton Dickinson) was used for data acquisition and the Flow Jo software was used for data analysis.

Immunofluorescence. Frozen sections (5 μm) were fixed in cold acetone, and then washed and blocked for 30 min with 10% (v/v) FCS in PBS. Sections were permeabilized in 0.1% (v/v) Tween in PBS for 10 min, and stained overnight at 4 °C with biotinylated anti-keratin 8 antibody (TROMA-1, Developmental Studies Hybridoma Bank). Sections were incubated with secondary antibody for 90 min at room temperature and, after washing three times for 15 min each in 0.1% (v/v) Tween in PBS, were blocked with anti-rat serum (Jackson Immuno Research Laboratories) for 30 min at room temperature. After washing, sections were incubated with anti-keratin 5 antibody (Covance) for 2 h at room temperature. Secondary reagents were streptavidin-Cy3 (Jackson Immuno Research Laboratories) and Alexa-Fluor-488-conjugated anti-rabbit antibody (Molecular Probes). Nuclei were counterstained with ProLong Gold antifade reagent with DAPI (Molecular Probes). Samples were analysed with an inverted confocal LSM 510 laser scanning confocal microscope (Zeiss).

GFP-LC3 analyses. GFP-LC3 mice and wild-type controls were maintained on a normal diet or kept without food for 48 h before analyses, with free access to drinking water. Before analyses, mice were anaesthetized and perfused through the left ventricle with 4% paraformaldehyde in PBS. Organs were collected and further fixed with the same fixative for 4 h, followed by treatment with 15% sucrose in PBS for 4 h at room temperature and then with 30% sucrose solution overnight at 4 °C. Tissue samples were embedded in OCT (Sakura Finetek Europe B.V.) and stored at -70 °C. The samples were sectioned (5 μm), and stained with Alexa-Fluor-488-conjugated anti-GFP antibody (Molecular Probes) overnight at 4 °C. Samples were mounted using ProLong Gold antifade reagent with DAPI.

Quantification of autophagic activity in thymic stromal cells. Thymi from two-week-old GFP-LC3 mice were cut into small pieces and digested at 37 °C in IMDM containing 0.2 mg ml⁻¹ collagenase (Roche), 0.2 mg ml⁻¹ dispase I (Roche), 2% FCS, 25 mM HEPES (pH 7.2) and 25 μg ml⁻¹ DNase I, followed by incubation in 5 mM EDTA for 5 min. Cells were washed and resuspended in Percoll (GE Healthcare), followed by a layer of Percoll and PBS as the upper phase. Gradients were spun for 30 min at 1,350g in the cold, and low density cells were collected from the upper interface, washed and then stained for FACS sorting. Stromal cells were sorted according to CD45, Ly51, EpCAM, CD80 and CD11c expression (cTECs: CD45⁺Ly51⁺EpCAM⁺; mTEC^{lo}: CD45⁺Ly51⁺EpCAM⁺CD80^{lo}; mTEC^{hi}: CD45⁺Ly51⁺EpCAM⁺CD80^{hi}; dendritic cells: CD45⁺CD11c⁺) and fixed on to poly-L-lysine-coated slides (Sigma Aldrich) with ice cold acetone. Cells were permeabilized in 0.1% (v/v) Tween in

PBS for 10 min, and blocked with 10% (v/v) FCS in PBS for 30 min at room temperature. Cells were incubated overnight at 4 °C with Alexa-Fluor-488-conjugated anti-GFP antibody and mounted with ProLong Gold antifade reagent with DAPI. A total of 1,000 dendritic cells, 1,845 cTECs, 1,232 mTEC^{hi} and 245 mTEC^{lo} have been analysed for determination of the percentage of autophagy-positive cells. Cells with more than five autophagosomes (that is, GFP-positive punctae between 0.5 and 1.5 μm in diameter) were scored as positive. For quantification of the number of autophagosomes per cell, a total of 81 autophagy-positive cTECs and 51 autophagy positive mTEC^{hi} have been recorded and analysed. Cells were analysed with an inverted confocal LSM510 Zeiss microscope by taking Z stacks of 0.5 μm thickness and subsequently the number of autophagosomes in the whole cell was determined.

Purification and adoptive transfer of T cells. Pooled cell suspensions of spleen and lymph nodes (mesenteric, inguinal, axillary, brachial, superficial cervical, deep cervical and lumbar) from *nu/nu*^{Atg5}^{-/-} and *nu/nu*^{WT} chimaeras were subjected to erythrocyte lysis. Cells were then incubated with Fc-receptor-blocking antibody (2.4G2) and were stained with biotin-anti-CD90.2 (30-H-12). After incubation with streptavidin microbeads (Miltenyi Biotec), CD90⁺ cells were positively selected on midi-MACS columns. 1×10^7 CD90⁺ cells were injected in a volume of 200 μl in PBS into the lateral tail veins of BALB/*c*^{nu/nu} animals. Animals were analysed 6–12 weeks after transfer.

Thymus transplantation. Embryonic thymi (E15; with or without deoxyguanosine treatment) were transplanted under the kidney capsule of female animals. *Atg5*^{-/-} embryos and *Atg5*^{+/+} (wild type) controls in individual experiments were obtained from the same *Atg5*^{+/+} pregnant female after mating to a *Atg5*^{+/+} male.

Deoxyguanosine treatment. E15 thymic lobes were placed on 0.45 mm membrane filters (Milipore) supported by Gelfoam (Pharmacia&Upjohn) and were incubated for 5 days with 10% (v/v) IMDM supplemented with 1.35 mM 2-deoxyguanosine before transplantation.

Bone marrow chimaeras. For mixed bone marrow chimaeras, T-cell-depleted TCR-HA *Rag2*^{-/-} CD45.1^{+/+} and DO11.10 *Rag2*^{-/-} CD45.1^{+/+} bone marrow were mixed at a ratio of 1:1 and a total of 1×10^7 cells were injected into the lateral tail vein of irradiated BALB/c recipients (2×450 rad). Animals were transplanted with *Atg5*^{-/-} or *Atg5*^{+/+} (wild-type) lobes 4 weeks after bone marrow reconstitution. Grafts were analysed 6 weeks after transplantation.

Fetal liver chimaeras. Fetal livers were harvested from E15 embryos. A total of 5.5×10^6 fetal liver cells were injected into the lateral tail veins of irradiated (2×550 rad) CD45.1 C57BL/6 recipients. Chimaeras were analysed 12 weeks after reconstitution.

Semiquantitative RT-PCR. mTECs were isolated from 4–5 transplanted thymi as described above and total RNA was isolated using the high pure RNA isolation kit (Roche). RNA was reverse transcribed using the iScript RT kit (Biorad). PCR reactions were carried out in a final volume of 25 μl using the Advantage 2 polymerase mix (Clontech) and Advantage 2 PCR buffer SA (Clontech) according to the manufacturers' recommendations. dNTPs (Fermentas) were used at a final concentration of 200 μM and primers were used each at 400 nM final concentration.

Primers were *Aire* forward: 5'-GAGTCACAGCACCTTCCTCTT-3' (438 base pairs (bp)); *Aire* reverse: 5'-GGGACAGCCGTCACAACA-3' (438 bp); *Spt1* forward: 5'-TGAAACTCAGGCAGATAG-3' (383 bp); *Spt1* reverse: 5'-GAGGGAG AATAGTCAGGAT-3' (383 bp); *Tjff3* forward: 5'-CTGGCTAATGCTGTGG TG-3' (377 bp); *Tjff3* reverse: 5'-TGTTGGCTGTGAGGTCTTT-3' (377 bp); *Csn2* forward: 5'-TCATCCTCGCCTGCCTTGT-3' (401 bp); *Csn2* reverse: 5'-GCCGAGCACAGTTTCAGATT-3' (401 bp); *Actb* forward: 5'-GGTGGG AATGGGTCAGA-3' (380 bp); *Actb* reverse: 5'-GAGCATAGCCCTCGTAG AT-3' (380 bp).

All amplicons span at least one intron. Cycling conditions were as follows: 94 °C for 3 min; $\times 31$ (*Actb*), $\times 34$ (*Aire*), $\times 36$ (*Csn2*/*Tjff3*), $\times 38$ (*Spt1*) 94 °C for 20 s, 56 °C for 30 s, 72 °C for 30 s; 72 °C for 5 min. Samples were run on a 1.5% (w/v) agarose and ethidium bromide gel in TAE buffer.

Histopathology. Organs were collected from donor animals and immersion-fixed with either 4% paraformaldehyde overnight at 4 °C or in an IHC zinc fixative (Becton Dickinson) for 48 h at room temperature. Samples were embedded in paraffin blocks. Sections (3–5 μm thick) were stained with haematoxylin and eosin and automatically scanned using the Zeiss MIRAX SCAN system. Subsequently, sections were analysed 'blind' using the Zeiss MIRAX VIEWER software.

Statistical analyses. All statistical analyses were performed using the two tailed Student's *t*-test with unequal variance.

31. Klein, L., Klein, T., Ruther, U. & Kyewski, B. CD4 T cell tolerance to human C-reactive protein, an inducible serum protein, is mediated by medullary thymic epithelium. *J. Exp. Med.* **188**, 5–16 (1998).
32. Kirberg, J. *et al.* Thymic selection of CD8⁺ single positive cells with a class II major histocompatibility complex-restricted receptor. *J. Exp. Med.* **180**, 25–34 (1994).

Regulatory networks define phenotypic classes of human stem cell lines

Franz-Josef Müller^{1,2}, Louise C. Laurent^{1,3}, Dennis Kostka^{4†}, Igor Ulitsky⁵, Roy Williams⁶, Christina Lu¹, In-Hyun Park⁷, Mahendra S. Rao^{8,9}, Ron Shamir⁵, Philip H. Schwartz^{10,11}, Nils O. Schmidt¹² & Jeanne F. Loring^{1,6}

Stem cells are defined as self-renewing cell populations that can differentiate into multiple distinct cell types. However, hundreds of different human cell lines from embryonic, fetal and adult sources have been called stem cells, even though they range from pluripotent cells—typified by embryonic stem cells, which are capable of virtually unlimited proliferation and differentiation—to adult stem cell lines, which can generate a far more limited repertoire of differentiated cell types. The rapid increase in reports of new sources of stem cells and their anticipated value to regenerative medicine^{1,2} has highlighted the need for a general, reproducible method for classification of these cells³. We report here the creation and analysis of a database of global gene expression profiles (which we call the ‘stem cell matrix’) that enables the classification of cultured human stem cells in the context of a wide variety of pluripotent, multipotent and differentiated cell types. Using an unsupervised clustering method^{4,5} to categorize a collection of ~150 cell samples, we discovered that pluripotent stem cell lines group together, whereas other cell types, including brain-derived neural stem cell lines, are very diverse. Using further bioinformatic analysis⁶ we uncovered a protein–protein network (PluriNet) that is shared by the pluripotent cells (embryonic stem cells, embryonal carcinomas and induced pluripotent cells). Analysis of published data showed that the PluriNet seems to be a common characteristic of pluripotent cells, including mouse embryonic stem and induced pluripotent cells and human oocytes. Our results offer a new strategy for classifying stem cells and support the idea that pluripotency and self-renewal are under tight control by specific molecular networks.

Cultured cell populations are traditionally classified as having the qualities of stem cells by their expression of immunocytochemical or PCR markers⁷. This approach can often be misleading if these markers are used to categorize novel stem cell preparations or predict inherent multipotent or pluripotent features⁸. To develop a more robust classification system, we created a framework for identifying putative novel stem cell preparations by their whole-genome messenger RNA expression phenotypes (Fig. 1). The core reference data set, which we call the ‘stem cell matrix’, includes cultures of human cells that have been reported to have either stem cell or progenitor qualities, including human embryonic stem cells, mesenchymal stem cells and neural stem cells. To provide the context in which to place the stem cells, we included non-stem-cell samples such as fibroblasts and differentiated embryonic stem cell derivatives. To avoid biasing

the classification methods, it was critical that we designated the input cell types with terminology that carried as little preconception about

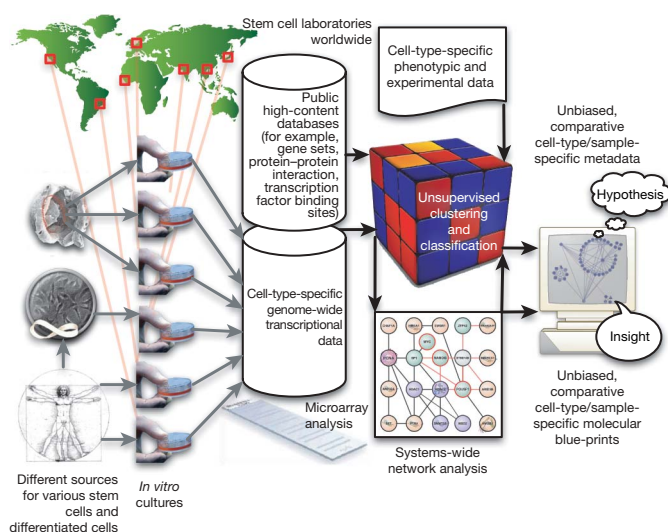


Figure 1 | Sample collection and analysis for the stem cell matrix. Cell preparations for the stem cell matrix are cultured in the authors' laboratories or collected from other sources worldwide. Samples are assigned source codes that capture their biological origin and a relatively unbiased description of the cell type (such as BNLin for brain-derived neural lineage). Samples are collected and processed at a central laboratory for microarray analysis on a single Illumina BeadStation instrument. The genomics data are processed by unsupervised algorithms that are capable of grouping the samples based on non-obvious expression patterns encoded in transcriptional phenotypes. For pathway discovery, existing high-content databases with experimental data (for example, protein–protein interaction data or gene sets) are combined with our transcriptional database, a priori assumed identity of cell types and bootstrapped sparse non-negative matrix factorization (sample clustering) to produce metadata that can be mined with GSA software and topology-based gene set discovery methods (systems-wide network analysis). Web-based, computer-aided visualization methodologies can be used by researchers to formulate testable hypotheses and generate results and insights in stem cell biology. Two exemplary results we report in this paper are the classification of novel stem cell types in the context of other better understood stem cell preparations, and a molecular map of interacting proteins that appear to function together in pluripotent stem cells.

¹Center for Regenerative Medicine, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA. ²Center for Psychiatry, ZIP-Kiel, University Hospital Schleswig Holstein, Niemannsweg 147, D-24105 Kiel, Germany. ³University of California, San Diego, Department of Reproductive Medicine, 200 West Arbor Drive, San Diego, California 92035, USA. ⁴Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Ihnestrasse 63-73, D-14195 Berlin, Germany. ⁵School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel. ⁶The Burnham Institute for Medical Research, 10901 North Torrey Pines Road, La Jolla, California 92037, USA. ⁷Division of Pediatric Hematology/Oncology, Children's Hospital Boston and Dana Farber Cancer Institute, Boston, Massachusetts 02115, USA. ⁸Invitrogen Co, 3705 Executive Way, Frederick, Maryland 21704, USA. ⁹Center for Stem Cell Biology, Buck Institute on Aging, 8001 Redwood Boulevard, Novato, California 94945, USA. ¹⁰Center for Neuroscience Research, Children's Hospital of Orange County Research Institute, 455 South Main Street, Orange, California 92868, USA. ¹¹Developmental Biology Center, University of California, Irvine, 4205 McGaugh Hall, Irvine, California 92697, USA. ¹²Department for Neurosurgery University Medical Center Hamburg-Eppendorf, Martinistrasse 52, D-20246 Hamburg, Germany. [†]Present address: Genome and Biomedical Sciences Facility and Department of Statistics, University of California, Davis 451 Health Sciences Drive, Davis, California 95616, USA.

their identity as possible. Our nomenclature ('source code') has two components: the first is the tissue or cultured cell line of origin. The second term captures a description of the culture itself. Supplementary Tables 1–8 summarize the descriptions of the core samples and their assigned source codes.

To sort the cell types we used an unsupervised machine learning approach to cluster transcriptional profiles of the cell preparations into stable distinct groups. Sparse non-negative matrix factorization (sNMF) was adjusted for this task by implementing a bootstrapping algorithm to find the most stable groupings (see also Supplementary Discussion 1)^{4,5}. The stability of the clustering⁹ indicated that the data set most likely contained about 12 different types of samples (Fig. 2a and Supplementary Methods 2). The composition of the stable clusters revealed both predictable and unpredicted groupings of a priori designations (Fig. 2b and Supplementary Fig. 1). The 20 samples identified as undifferentiated human pluripotent stem cell (PSC) preparations were grouped together in one dominant cluster (Fig. 2, cluster 1) and one secondary cluster (Fig. 2, cluster 5). Sixty-two of the samples were brain-derived cells that were described as neural stem or progenitor cells based on their source, culture methods and classical markers. Most of the designated neural stem cells were distributed among multiple clusters, indicating a great deal of diversity in neural stem cell preparations. But one group of the brain-derived lines, those derived from surgical specimens from living patients (HANSE cells, see below), remained together throughout the iterative clusterings (Fig. 2, cluster 6; see also Supplementary Fig. 3 and Supplementary Methods 1). The HANSE cell group consisted of transcriptional profiles that were derived from neurosurgical specimens following published protocols for multipotent neural progenitor derivation and propagation^{10,11}. These cells expressed markers that are commonly used to identify neural stem cells¹² (see Supplementary Fig. 4), but the clustering clearly separated them from the other samples that had been derived from post-mortem brains of prematurely born infants (SC23 and SC30, see Fig. 2b)^{10,11}.

We tested the ability of our data set to categorize additional preparations by adding 66 samples comprising new cultures derived from PSC lines that were already in the matrix, preparations that were not yet included (but their presumptive cell type was already represented), or new cell types. We chose two new types of cells: a differentiated cell type (umbilical vein endothelial cells (HUVECs)) and a recently developed new source of pluripotent cells called induced pluripotent stem cells^{13–16} (iPSCs, Supplementary Table 9). iPSCs have been generated from somatic cells, including adult fibroblasts, by genetic manipulation of certain transcription factors^{13,15–17}. We re-computed clustering results including the test data set (Supplementary Table 10). All of the HUVEC samples clustered together and formed a distinct group. Most of the additional PSC lines (human embryonic stem cells (embryonic PSCs; ePSCs) and iPSCs) from several different laboratories were placed into a context that contained solely PSC lines. Three additional germ cell tumour lines clustered together with the tumour-derived pluripotent stem cell (tPSC) line 2102Ep and samples of three human embryonic stem (ES) cell lines: BG01v (ref. 18), Hues7 (ref. 19) and Hues13 (ref. 19). BG01v is an established aneuploid variant line and the two Hues lines are aneuploid variants of the originally euploid lines (not shown).

We used a combination of analysis tools to explore the basis of the unsupervised classification of the samples in the core data set. Gene Set Analysis²⁰ (GSA) is a means to identify the underlying themes in transcriptional data in terms of their biological relevance.

GSA uses lists of genes²⁰ that are related in some way; the common criterion is that the relationships among the genes in the lists are supported by empirical evidence²⁰. GSA highlighted numerous significant differences among the computationally defined categories. (See Supplementary Fig. 2, Supplementary Table 11, Supplementary Methods and <http://www.stemcellmatrix.org>).

Although GSA is valuable for discovering specific differences among sample groups, it is limited to curated gene lists and cannot be used to

discover new regulatory networks. The MATISSE algorithm⁶ (<http://acgt.cs.tau.ac.il/matisse>) takes predefined protein–protein interactions (for example, from yeast two-hybrid screens) and seeks connected subnetworks that manifest high similarity in sample subsets. The modified version used in this analysis is capable of extracting subnetworks that are co-expressed in many samples but also significantly upregulated or downregulated in a specific sample cluster.

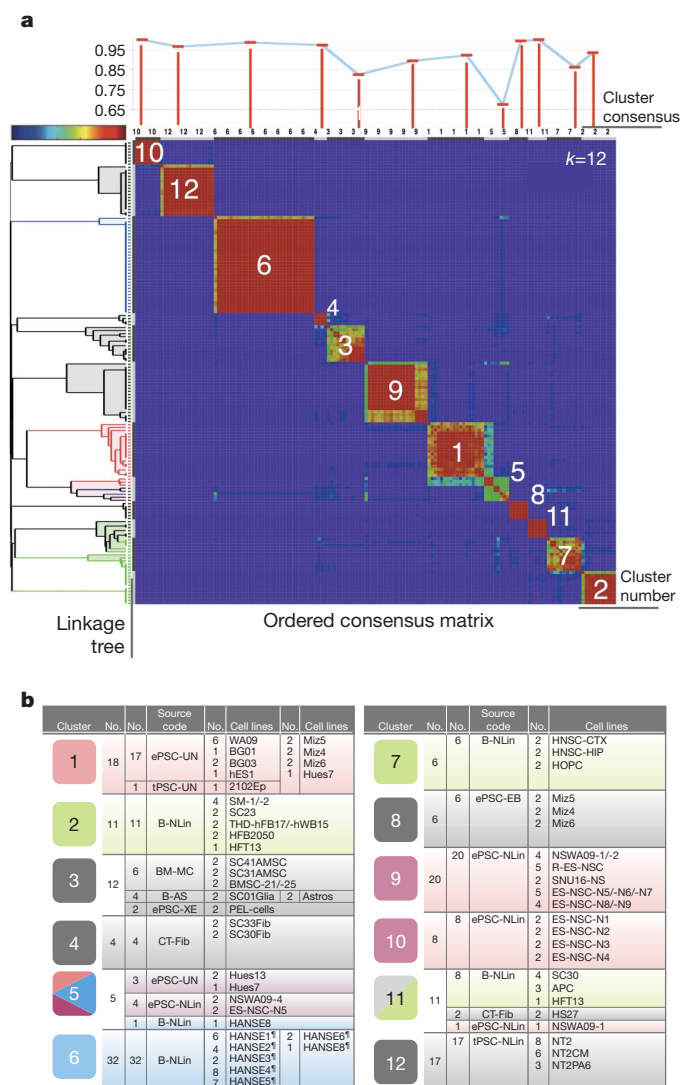


Figure 2 | Clusters of samples based on machine learning algorithm. Samples were distributed on the basis of their transcriptional profiles into consensus clusters using sNMF. **a**, Consensus matrix from consensus clustering results (centre matrix plot). The consensus matrix is a visual representation of the clustering results and the separation of the sample clusters from each other. Blue indicates no consensus; red indicates very high consensus. The numbers (1–12) on the diagonal row of clusters indicate the number assigned to the cluster by sNMF. These numbers (cluster 1 to cluster 12) are used throughout the text to indicate the group of samples in that cluster. The bar graph above the consensus matrix plot shows the summary statistics assessing the overall quality of each cluster. The cluster consensus value (0–1) is plotted above the corresponding cluster in the matrix plot. Note that most clusters (clusters 10, 12, 6, 4, 9, 1, 8, 11, 7 and 2) have a high-quality measurement. To the left of the consensus matrix is another view of the consensus data, visualized as a dendrogram. This is a representation of the hierarchical clustering tree of the consensus matrix. **b**, The content of the sample clusters resulting from the same sNMF run are displayed. Numbers are the same cluster numbers assigned by the consensus clustering algorithm that are used throughout the text and figures. For more information on samples, source code and references see Supplementary Tables 1–10. No., number of samples. The symbol ‘*’ indicates that samples were derived from adult brain specimens.

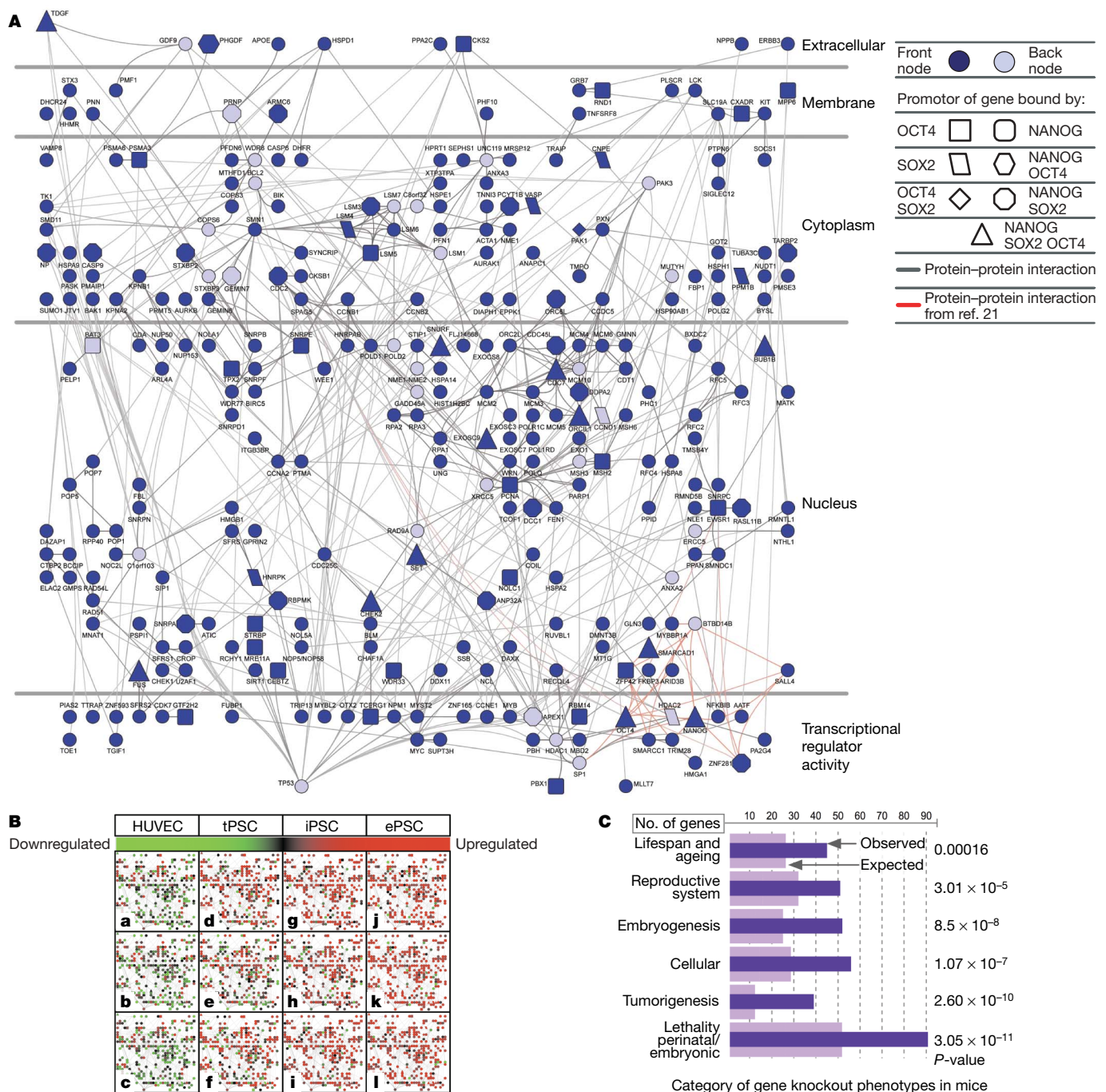


Figure 3 | Pluripotent stem-cell-specific protein-protein interaction network detected by MATISSE. Clusters from the sNMF $k = 12$ analysis were used in combination with the transcriptional database to identify protein-protein interaction networks enhanced in PSCs. **A**, A large differentially expressed connected subnetwork (PluriNet) shows the dominance of cell cycle regulatory networks in PSCs (see legend). All of the dark blue symbols are genes that are highly expressed in most PSCs compared to the other cell samples in the data set. Front nodes, as represented by stem cell matrix expression data, and back nodes, as inferred by MATISSE, are displayed with different colour shades⁶. Highlighted in red are the interactions of a group of proteins associated with pluripotency in murine ePSCs²¹. This subnetwork shows a significant enrichment in genes that are targeted in the genome by the transcription factors NANOG ($P = 5.88 \times 10^{-4}$), SOX2 ($P = 0.058$) and E2F ($P = 1.29 \times 10^{-16}$, all P -values are Bonferroni corrected). For an interactive visualization of PluriNet, see <http://www.stemcellmatrix.org>. **B**, Heat-map-like visualization of PluriNet genes for samples from the test data set: HUVECs (UC-EC, **a-c**), derived from three independent individuals), germ cell tumour-derived pluripotent stem cells (tPSC-UN, **d-f**, lines GCT-C4, GCT-72, GCT-27X,

derived from three independent individuals), induced pluripotent stem cells (iPSC-UN, **g-i**, BJ1-iPS12, MSC-iPS1, hFib2-iPS5, three independently derived lines from different somatic sources) and embryonic stem cells (ePSC-UN, **j-l**, lines Hues22, HSF6, ES2, derived from three independent blastocysts in three independent laboratories). Most PluriNet genes are markedly upregulated in iPSC-UN and ePSC-UN cells. tPSC-UN cells show a less consistent expression pattern. UC-EC cells show lower expression levels of most PluriNet genes. See Supplementary Fig. 5 for a larger version of the same heat maps. **C**, Analysis of genes from PluriNet in the context of phenotypes that have been reported to result from specific genetic manipulations (for example, gene knockout) in mice in the MGI 3.6 phenotype ontology database (<http://www.informatics.jax.org/>). We find significant over-representation of phenotypes 'lethality (perinatal/embryonic)', 'tumorigenesis', 'cellular', 'embryogenesis', 'reproductive system' and 'lifespan and ageing' among the genes in PluriNet. Although these broad categories might be rather unspecific surrogate markers for PSC function in mammals, this analysis might point towards PluriNet's role *in vivo*. For more details, see also Supplementary Fig. 6 and Supplementary Table 12.

Table 1 | PluriNet expression patterns in various model systems for pluripotency

a Expression of PluriNet genes in murine model systems			
Cell type	Upregulated/downregulated		
MII oocytes	Upregulated*		
Zygote	Upregulated*		
Embryo (two-cell blastocyst)	Upregulated*		
ePSC	Upregulated†		
EpiSC	Upregulated†		
iPSC	Upregulated†		
Fibroblasts (normal)	Downregulated‡		
Fibroblasts (transformed)	Downregulated‡		
b Successful PluriNet-based, post-hoc classification in murine model systems			
Cell type	Upregulated/downregulated	Pluripotency (PAM)	Germline transmission (PAM)
ePSC	Upregulated	Yes‡	Yes‡
EpiSC	Upregulated	Yes‡	Yes‡
iPSC	Upregulated	Yes‡	Yes‡
Fibroblasts (normal)	Downregulated	Yes‡	Yes‡
Fibroblasts (transformed)	Downregulated	Yes‡	Yes‡
c Expression of PluriNet genes in human model systems			
Cell type	Upregulated/downregulated		
MII oocytes	Upregulated§		
tPSC	Upregulated		
ePSC	Upregulated ¶		
iPSC	Upregulated ¶		
ePSC-derived cell types	Downregulated		
Somatic cell types	Downregulated ¶		
Somatic cancer cell line (HeLa)	Downregulated#		
d Successful PluriNet-based, post-hoc classification in human model systems			
Cell type	Upregulated/downregulated	Pluripotency (PAM)	
tPSC	Upregulated	Yes**	
ePSC	Upregulated	Yes**	
iPSC	Upregulated	Yes**	
ePSC-derived cell types	Downregulated	Yes**	
Somatic cell types	Downregulated	Yes**	

This table summarizes the expression patterns of PluriNet in various model systems of pluripotency and differentiation. More details on the specific tests and explanations of the data sources for the results can be found as indicated below. EpiSC, epiblast-derived stem cells²⁴; PAM, prediction analysis of microarray, classifier with leave-one-out cross validation²⁷. 'Yes' in parts **b** and **d** indicates correct classification of pluripotent state (pluripotent or not pluripotent) in >90% of samples.

* For more details see Supplementary Figs 8 and 9.

† For more details see Supplementary Fig. 10.

‡ For more details see Supplementary Fig. 10.

§ For more details see Supplementary Fig. 7.

|| For more details see Fig. 3B and Supplementary Figs 5 and 12.

¶ For more details see Supplementary Fig. 11.

For more details see Supplementary Discussion 2.

** For more details see Supplementary Fig. 12.

Because the PSC preparations were consistently clustered together we used MATISSE to look for distinctive molecular networks that might be associated with the unique PSC qualities of pluripotency and self-renewal. A Nanog-associated regulatory network has been outlined in mouse embryonic PSCs²¹, and we looked for the elements of this network in human PSCs using our unbiased algorithm. We found that the algorithm predicts that human PSCs possess a similar NANOG-linked network (Fig. 3A; elements labelled in red). However, we also discovered that the human NANOG network seems to be integrated as a small component of a much larger protein–protein interaction network that is upregulated in human PSCs (Fig. 3). Notably, this PSC-specific network (termed pluripotency-associated network, PluriNet) contains key regulators that are involved in the control of cell cycle, DNA replication, DNA repair, DNA methylation, SUMOylation, RNA processing, histone modification and nucleosome positioning (see also Supplementary Discussion 2 and <http://www.openstemcellwiki.org>). Many of the genes in the PluriNet have been linked to embryogenesis, tumorigenesis and ageing (Fig. 3C and Supplementary Fig. 6). We further explored the

hypothesis that pluripotency is closely linked to PluriNet expression by analysing published gene expression data sets from human oocytes, various types of PSCs and murine embryos (see Table 1 for a summary of our findings in various model systems). Analysis of a microarray data set²² that spans development from murine oocytes to the late blastocyst stage revealed that the PluriNet expression is dynamic and upregulated during early mammalian embryogenesis (Table 1 and Supplementary Figs 7–9)²³. Also, our preliminary analyses indicate that the PluriNet is strongly upregulated in mouse PSCs, mouse iPSCs and mouse epiblast-derived stem cells²⁴ when compared to somatic cells. Therefore the PluriNet may be useful as a biologically inspired gauge for classifying both murine and human PSC phenotypes (Table 1 and Supplementary Figs 10–13).

Our data indicate that an unbiased global molecular profiling approach combined with a transcriptional phenotype collection using suitable machine learning algorithms can be used to understand and codify the phenotypes of stem cells^{4,5,25}. Although it is more extensive than any stem cell data set reported so far, we consider our database and the PluriNet to be a work in progress. As more direct evidence for protein–protein interactions in human cells becomes available, it will be possible to refine the networks we have defined and make them more useful for testing hypotheses about the nature of stem cell pluripotency and multipotency. Also, our sample collection is limited to pluri- and multipotent stem cell types that grow well in culture, and does not include some of the most well studied lineages, such as haematopoietic stem cells. Resolution and reliability of a context-based unsupervised classification can be expected to grow with the breadth and depth of the database content²⁶. Even with these limitations, we have shown that the data set and PluriNet have already proved useful for categorizing cell types using unbiased criteria. As more stem cell populations become available, cultured by new methods, isolated from new sources, or induced by new methods, we will use the PluriNet and the stem cell matrix as a reference system for phenotyping the cells and comparing them with existing cell lines.

METHODS SUMMARY

For an overview of the general workflow, please also refer to Fig. 1. A detailed list of the samples, culture methods and reference publications is provided in Supplementary Information¹¹. Generally, RNA from each sample was prepared from approximately 1×10^6 cultured cells. Sample amplification, labelling and hybridization on Illumina WG8 and WG6 Sentrix BeadChips were performed for all arrays in this study according to the manufacturer's instructions (<http://www.illumina.com>) at a single Illumina BeadStation facility. We used the Consensus Clustering framework⁹ to cluster transcription profiles and to assess stability of the results. As the algorithm, we used sparse non-negative matrix factorization⁵. For data perturbation, 30 subsampling runs were performed for each considered number of clusters (k). In each run, 80% of the data was subjected to ten random restarts. The R-script can be downloaded at <http://www.stemcellmatrix.org>. Details on the application of GSA²⁰, PAM²⁷, MATISSE⁶ as well as publicly available data sets used in this study can be found in the Methods section. We modified the MATISSE⁶ computational framework to fit the goals of this study. For the present analysis we used the human physical interaction network that we had previously assembled⁶ and augmented it with additional interactions from recent publications^{21,28,29}. The 64 interactions in ref. 21 were mapped to the corresponding human orthologues using the NCBI HomoloGene database.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 15 December 2007; accepted 26 June 2008.

Published online 24 August 2008.

- Müller, F. J., Snyder, E. Y. & Loring, J. F. Gene therapy: can neural stem cells deliver? *Nature Rev. Neurosci.* **7**, 75–84 (2006).
- Murry, C. E. & Keller, G. Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* **132**, 661–680 (2008).
- Adewumi, O. *et al.* Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nature Biotechnol.* **25**, 803–816 (2007).
- Brunet, J. P., Tamayo, P., Golub, T. R. & Mesirov, J. P. Metagenes and molecular pattern discovery using matrix factorization. *Proc. Natl Acad. Sci. USA* **101**, 4164–4169 (2004).

5. Gao, Y. & Church, G. Improving molecular cancer class discovery through sparse non-negative matrix factorization. *Bioinformatics* **21**, 3970–3975 (2005).
6. Ulitsky, I. & Shamir, R. Identification of functional modules using network topology and high-throughput data. *BMC Syst. Biol.* **1**, 8 (2007).
7. Carpenter, M. K., Rosler, E. & Rao, M. S. Characterization and differentiation of human embryonic stem cells. *Cloning Stem Cells* **5**, 79–88 (2003).
8. Goldman, B. Magic marker myths. *Nature Reports Stem Cells*. doi:10.1038/stemcells.2008.26 (2008).
9. Monti, S., Tamayo, P., Mesirov, J. & Golub, T. Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Mach. Learn.* **52**, 91–118 (2003).
10. Palmer, T. D. *et al.* Cell culture. Progenitor cells from human brain after death. *Nature* **411**, 42–43 (2001).
11. Schwartz, P. H. *et al.* Isolation and characterization of neural progenitor cells from post-mortem human cortex. *J. Neurosci. Res.* **74**, 838–851 (2003).
12. Kornblum, H. I. & Geschwind, D. H. Molecular markers in CNS stem cell research: hitting a moving target. *Nature Rev. Neurosci.* **2**, 843–846 (2001).
13. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–676 (2006).
14. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–872 (2007).
15. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
16. Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
17. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
18. Zeng, X. *et al.* BG01V: a variant human embryonic stem cell line which exhibits rapid growth after passaging and reliable dopaminergic differentiation. *Restor. Neurol. Neurosci.* **22**, 421–428 (2004).
19. Cowan, C. A. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *N. Engl. J. Med.* **350**, 1353–1356 (2004).
20. Efron, B. & Tibshirani, R. On testing the significance of sets of genes. *Ann. Appl. Stat.* **1**, 107–129 (2007).
21. Wang, J. *et al.* A protein interaction network for pluripotency of embryonic stem cells. *Nature* **444**, 364–368 (2006).
22. Wang, Q. T. *et al.* A genome-wide study of gene activity reveals developmental signaling pathways in the preimplantation mouse embryo. *Dev. Cell* **6**, 133–144 (2004).
23. Chambers, I. *et al.* Nanog safeguards pluripotency and mediates germline development. *Nature* **450**, 1230–1234 (2007).
24. Tesar, P. J. *et al.* New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* **448**, 196–199 (2007).
25. Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
26. Donoho, D. & Stodden, V. When does non-negative matrix factorization give correct decomposition into parts? *Proc. NIPS* (2003) (http://books.nips.cc/papers/files/nips16/NIPS2003_LT10.ps.gz).
27. Lacayo, N. J. *et al.* Gene expression profiles at diagnosis in *de novo* childhood AML patients identify FLT3 mutations with good clinical outcomes. *Blood* **104**, 2646–2654 (2004).
28. Ewing, R. M. *et al.* Large-scale mapping of human protein-protein interactions by mass spectrometry. *Mol. Syst. Biol.* **3**, 89 (2007).
29. Mishra, G. R. *et al.* Human protein reference database–2006 update. *Nucleic Acids Res.* **34**, D411–D414 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. Stubban, H. Dittmer, S. Zapf and H. Meissner for their work with various cell cultures. We are grateful to D. Wakeman, R. Gonzalez, S. McKercher, J. P. Lee, H.-S. Park and S. Y. Moon for sharing their cell preparations for the type collection. We are also grateful to R. Wesselschmidt and M. Pera for their unique GCT lines and G. Daley for providing human iPSCs. A. M. Kocabas and J. Cibelli shared their human oocyte expression data with us. A. Barsky let us use the Cerebral 2.0 plug-in before its publication. M. Rosentraeger helped to compile the cell culture metadata. We thank J. Aldenhoff, D. Hinze-Selch, M. Westphal, K. Lamszus, U. Kehler, D. Barker and A. Fritz for their support and discussions of this project. This study has been supported by the following grants and awards: Christian-Abrechts University Young Investigator Award (F.-J.M.), SFB-654/C5 Sleep and Plasticity (F.-J.M. and D. Hinze-Selch), Hamburger Krebsgesellschaft Grant (N.O.S.), Edmond J. Safra Bioinformatics program fellowship at Tel-Aviv University (I.U.), Converging Technologies Program of The Israel Science Foundation Grant No 1767.07 (R.S.), Raymond and Beverly Sackler Chair in Bioinformatics (R.S.), Reproductive Scientist Development Program Scholar Award K12 5K12HD000849-20 (L.C.L.), California Institute for Regenerative Medicine Clinical Scholar Award (L.C.L.), NIH P20 GM075059-01 (J.F.L.), the Alzheimer's Association (J.F.L.), and anonymous donations in support of stem cell research.

Author Contributions J.F.L. and F.-J.M. designed the study and wrote the manuscript; I.U., R.W., D.K., R.S., L.C.L. and F.-J.M. designed and conducted the bioinformatics analysis; L.C.L., C.L., P.H.S., M.S.R., I.-H.P., F.-J.M. and N.O.S. conducted experiments and provided essential materials for this study.

Author Information The microarray data have been deposited at NCBI GEO (accession number GSE11508) and can also be accessed, processed and downloaded at <http://www.stemcellmesa.org>. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.-J.M. (fj.mueller@zip-kiel.de) or J.F.L. (jloring@scripps.edu).

METHODS

Compilation of type collection. Samples were either grown in our own laboratory or provided by collaborators. Each sample was prepared from approximately 1×10^6 cultured cells, which were mechanically harvested, pelleted and snap frozen in liquid nitrogen. Biological replicates were produced for almost all samples. Details on the included cell lines and culture methods can be found in the Supplementary Tables 3–8.

Neural progenitor cultures (HANSE) from neurosurgical specimens. Brain tissue samples were obtained from patients who underwent surgery for intractable temporal lobe epilepsy at the Department of Neurosurgery, University Medical Center Hamburg-Eppendorf, Germany ($n = 6$; 4 males and 2 females; mean age 33). All procedures were performed with informed consent and in accordance with institutional human tissue handling guidelines. We used modifications of reported protocols for establishing neural progenitor cultures from fetal and postmortem brain tissue^{10,30}. A more detailed description can be found in Supplementary Methods 1.

Whole-genome gene expression. All RNA was purified in our laboratory using standard methods. Sample amplification, labelling and hybridization on Illumina WG8 and WG6 Sentrix BeadChips were performed for all arrays in this study according to the manufacturer's instructions (Illumina) using an Illumina BeadStation (Burnham Institute Microarray Core).

Microarray data pre-processing. Raw data extraction was performed with BeadStudio v1.5 and probes with a detection score of less than 0.99 in all of the samples were discarded. The resulting probes were then quantile-normalized to correct for between-sample variation³¹. The sample data were quality controlled before normalization using the quality parameters provided by BeadStudio software. Before and after normalization the arrays were inspected with signal distribution box plots and by using the maCorrPlot package³².

Parameters for unsupervised classification. The data sets and the sparseness factor λ were adjusted for the unsupervised clustering task following previous reports^{4,5}. Parameters we have used for this study are: SCM core data set (153 samples), $\lambda = 0.01$; SCM test data set (219 samples), $\lambda = 0.1$. The pre-processed data sets used can be downloaded at <http://www.stemcellmatrix.info>.

Gene expression and gene set analysis. To screen for differentially expressed groups of genes between computationally defined sample clusters we used the Gene Set Analysis (GSA) methods proposed previously^{33,34}. GSA was chosen because it uses a stringent max-mean algorithm to identify significantly differentially regulated gene sets. The cutoff P -value was adjusted to accommodate a false discovery rate (FDR) of 10%. A translation file was built to use GSA with Illumina expression data. We collected gene lists from recent publications and public repositories (MolSigDB2, Stanford repository). These files can be downloaded from <http://www.stemcellmatrix.org>. To screen for differentially expressed genes between computationally defined sample clusters we used standard t -test-based methods implemented in the R Bioconductor package³⁵. The cutoff P -value was adjusted to accommodate a FDR of 5%.

Detection of cluster-specific subnetworks using MATISSE. MATISSE⁶ (<http://acgt.cs.tau.ac.il/matisse>) was adjusted to detect differentially expressed connected subnetworks (DECSs), corresponding to connected subnetworks in a physical interaction network that show a significant co-expression pattern. The physical network used by MATISSE contains vertices corresponding to genes and edges corresponding to protein–protein and protein–DNA interactions. For the present analysis we used the human physical interaction network that we had previously assembled⁶ and augmented it with additional interactions from recent publications^{21,28,29}. In total, the network contained 34,212 interactions among 9,355 proteins.

Originally, MATISSE used the Pearson correlation coefficient as a measure of similarity between the expression patterns of gene pairs. These similarity values serve as a starting point for the computation of pair-wise weights using a probabilistic model. The Pearson correlation between a pair of genes captures a global similarity trend between their expression patterns. In this work we were interested in extracting groups of genes that are not only similar across the experimental conditions, but also show significantly high or significantly low expression values in a specific subset of the samples, identified using the sNMF clustering scheme. To this end we devised a hybrid similarity score that captures two features: (1) both genes show differential expression; (2) the genes have similar expression patterns, regardless of their differential expression.

We denote the expression pattern of gene i by $x^i = (x^i_1, x^i_2, \dots, x^i_m)$. Assume that we are interested in DECSs upregulated in a condition subset $A \subseteq \{1, \dots, m\}$. To address goal (1), we use an 'ideal' expression profile $p = (p_1, p_2, \dots, p_m)$ where $p_i = 1$ if $i \in A$ and $p_i = -1$ otherwise. The signs are reversed if we are interested in a DECS downregulated in A . r_{kp} is the Pearson correlation coefficient between x^k and p . Intuitively, r_{kp} is close to 1 if the corresponding transcript is strongly upregulated in A compared to the other conditions, and close to -1 if it is

strongly downregulated in A . This measure has been suggested as an aparametric differential expression score³⁶. Note that the Pearson correlation is invariant under normalization of the patterns to zero mean and standard deviation of 1. For every gene pair (i, j) we compute $S_{\text{diff}}(i, j) = (r_{ip} + r_{jp})/2$. To address goal (2) we use the partial correlation coefficient between the gene patterns conditioned on the ideal profile. Formally, $S_{\text{part}}(i, j) = \frac{r_{x^i, x^j} - r_{x^i, p} r_{x^j, p}}{\sqrt{(1 - r_{x^i, p}^2)(1 - r_{x^j, p}^2)}}$, where r_{yz} is the

Pearson correlation coefficient between the profiles y and z . Intuitively, S_{part} conveys the information about how similar x^i and x^j are, regardless of their differential expression in A . Finally, we use the similarity score $S = \lambda S_{\text{diff}} + S_{\text{part}}$, where λ is a trade-off parameter setting the relative importance of the differential expression in the similarity score. We used $\lambda = 3$ for the analysis described in this paper. These S scores are then modelled using the probabilistic model described previously⁶. The advantage of using this pair-wise scoring scheme over the use of gene-specific differential expression scores, such as those proposed by others³⁷, is that it will prefer gene groups that are not only differentially expressed in the specified condition subset, but also have coherent expression profiles.

To diminish the effect of the size difference between the clusters, we reduced the number of conditions in clusters 1, 2, 3, 6, 9, 10 and 12, by including fewer replicates. Overall, 105 samples were used in the MATISSE analysis and can be downloaded at <http://www.stemcellmatrix.org>. We executed this MATISSE variant iteratively, each time setting A to contain all the samples of a single cluster or a cluster pair. The upper bound on module size was set to 300 and the rest of the parameters were as previously reported⁶. We then filtered the resulting networks by removing DECSs that overlapped more than 50% with other, higher scoring DECSs. The full set of the DECSs is available at <http://www.stemcellmatrix.org>.

Visualization. For visualization of the selected DECS we used Cytoscape 2.5 (ref. 38) and Cerebral 2.0 (ref. 39). Localization data from HRPD and the GO-Molecular function categories were also used²⁹. NANOG, POU5F1/OCT4 and SOX2 promoter binding information was used to code the ESC-specific regulation of nodes⁴⁰. Permutmatrix was used for heat maps⁴¹. Data for the analysis of human oocytes were accessed on the authors' or the journals' website⁴². For analysis of iPSCs induced with LIN28, OCT4, NANOG and SOX2, the data set was obtained from the Thomson laboratory¹⁵.

Classification based on PluriNet. We used the 299 genes from DECS (Up(1,5)A) (PluriNet) with the PAM²⁰ software package. Class probabilities were re-computed 10,000 times; average scores are reported in Supplementary Figs 10 and 12. We translated the human genes into their murine orthologues from PluriNet using the NCBI HomoloGene database for re-analysing murine expression profiles. The expression array data from murine fibroblasts, induced pluripotent cells, epiblast-derived stem cells and murine embryonic stem cells were downloaded from NCBI GEO^{21–24}.

30. Imitola, J. *et al.* Directed migration of neural stem cells to sites of CNS injury by the stromal cell-derived factor 1 α /CXCR4 chemokine receptor 4 pathway. *Proc. Natl Acad. Sci. USA* **101**, 18117–18122 (2004).
31. Barnes, M., Freudenberg, J., Thompson, S., Aronow, B. & Pavlidis, P. Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms. *Nucleic Acids Res.* **33**, 5914–5923 (2005).
32. Ploner, A., Miller, L. D., Hall, P., Bergh, J. & Pawitan, Y. Correlation test to assess low-level processing of high-density oligonucleotide microarray data. *BMC Bioinformatics* **6**, 80 (2005).
33. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
34. Efron, B. & Tibshirani, R. On testing the significance of sets of genes. *Ann. Appl. Stat.* **1**, 107–129 (2007).
35. R Development Core Team, R. A language and environment for statistical computing, help files. (<http://www.bioconductor.org/>) (2007).
36. Troyanskaya, O., Garber, M., Brown, P., Botstein, D. & Altman, R. Nonparametric methods for identifying differentially expressed genes in microarray data. *Bioinformatics* **18**, 1454–1461 (2002).
37. Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A. F. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* **18** (suppl. 1), S233–S240 (2002).
38. Cline, M. S. *et al.* Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2**, 2366–2382 (2007).
39. Barsky, A., Gardy, J. L., Hancock, R. E. & Munzner, T. Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation. *Bioinformatics* **23**, 1040–1042 (2007).
40. Boyer, L. A. *et al.* Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947–956 (2005).
41. Caraux, G. & Pinloche, S. PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* **21**, 1280–1281 (2005).
42. Kocbas, A. *et al.* The transcriptome of human oocytes. *Proc. Natl Acad. Sci. USA* **103**, 14027–14032 (2006).

LETTERS

A paracrine requirement for hedgehog signalling in cancer

Robert L. Yauch^{1*}, Stephen E. Gould^{1*}, Suzie J. Scales¹, Tracy Tang¹, Hua Tian¹, Christina P. Ahn¹, Derek Marshall¹, Ling Fu¹, Thomas Januario¹, Dara Kallop¹, Michelle Nannini-Pepe¹, Karen Kotkow^{2†}, James C. Marsters Jr¹, Lee L. Rubin^{2†} & Frederic J. de Sauvage¹

Ligand-dependent activation of the hedgehog (Hh) signalling pathway has been associated with tumorigenesis in a number of human tissues^{1–7}. Here we show that, although previous reports have described a cell-autonomous role for Hh signalling in these tumours^{1–7}, Hh ligands fail to activate signalling in tumour epithelial cells. In contrast, our data support ligand-dependent activation of the Hh pathway in the stromal microenvironment. Specific inhibition of Hh signalling using small molecule inhibitors, a neutralizing anti-Hh antibody or genetic deletion of *smoothed (Smo)* in the mouse stroma results in growth inhibition in xenograft tumour models. Taken together, these studies demonstrate a paracrine requirement for Hh ligand signalling in the tumorigenesis of Hh-expressing cancers and have important implications for the development of Hh pathway antagonists in cancer.

An increasing number of solid tumours have been reported to be reliant on ligand-dependent Hh pathway signalling within the tumour epithelium^{1–7}. Because the typical role for Hh in development is mediated by paracrine effects on mesenchymal cells^{8,9}, we wanted to explore further the mechanisms behind Hh-dependent tumorigenesis. We have evaluated the effect of a specific and potent small molecule SMO-acting antagonist of the Hh pathway (HhAntag) on cell viability across a large panel of cancer cell lines (Fig. 1a)^{10,11}. HhAntag has been demonstrated to be ~10-times more potent than the natural product SMO antagonist, cyclopamine, at inhibiting Hh pathway activity¹⁰. We observed a range of cellular sensitivities to HhAntag with half-maximal inhibitory concentration (IC₅₀) values for growth inhibition ranging from ~2 µM to >30 µM. In contrast to previous reports^{6,7}, no tissue specificity of *in vitro* sensitivity to HhAntag was observed. The extent of basal Hh pathway activity in these cell lines, as measured by the expression of *GLI1*, a well characterized endogenous direct transcriptional target gene, did not correlate with cellular sensitivity to HhAntag (Fig. 1a), nor could we detect a positive correlation between cellular sensitivity and other Hh pathway genes (Supplementary Fig. 1 and Supplementary Table 1). Finally, these differences were not a unique attribute of the synthetic small molecule HhAntag, as a similar discordance of all variables was observed with cyclopamine (Supplementary Fig. 2)¹².

The requirement for high concentrations of HhAntag or cyclopamine to elicit growth inhibition and the lack of correlation with Hh pathway activation suggested that the observed *in vitro* growth repression might be due to non-specific effects of these molecules. This is supported by the observation that only 100 nM HhAntag is needed to completely inhibit Hh signalling in a Hh-responsive human mesenchymal cell line (HEPM) expressing a *GLI* luciferase

reporter construct (HEPM-rep, Fig. 1b) and the IC₅₀ of 5 nM is ~400-times lower than that required to inhibit cell growth by 50% in the most sensitive cancer cell line (1.9 µM). To address specifically whether the *in vitro* growth inhibitory effect of high concentrations of Hh antagonists is due to off-target activity, we focused on two pancreatic cell lines described recently: CFPAC-1, a SMO-positive and cyclopamine-sensitive cell line, and BxPC-3, a SMO-negative cell line with differentially reported sensitivity to cyclopamine^{4,6,13}. Despite differences in SMO expression (Supplementary Table 1), we found that these cell lines exhibit similar sensitivities to the growth inhibitory effects of high concentrations of cyclopamine and HhAntag (Fig. 1a, b). Conversely, no growth inhibition was observed using two completely different mechanisms of Hh pathway inhibition: PKA activation using the agonist forskolin¹⁴ (Fig. 1c) or the anti-Hh ligand blocking antibody 5E1 (Fig. 1d). Finally, neither Hh pathway activation with recombinant Sonic hedgehog homologue (rSHH) nor pathway inhibition with high concentrations (10 µM) of Hh antagonists had any effect on endogenous *GLI1* messenger RNA levels in these cells (Fig. 1e), despite regulating pathway activity in HEPM cells. Previous reports relied on transfected *GLI* reporter constructs to monitor endogenous Hh pathway activity. rSHH was unable to activate *GLI* reporter activity when transfected into either CFPAC-1 or BxPC-3 cells (Supplementary Fig. 3). Although high concentrations of HhAntag and cyclopamine inhibited *GLI* reporter activity in CFPAC-1 and BxPC-3 cells, comparable repressions of unrelated transcriptional reporters were also observed at these high concentrations (Supplementary Fig. 4). Similar effects of Hh antagonists on viability and signalling were observed across a variety of cell lines that have been used previously (Supplementary Fig. 5)^{5–7}. Taken together, the data strongly argue against Hh signalling in these epithelial cancer cell lines and suggest that the previously observed effects of Hh antagonists on growth and reporter activity are due to off-target effects of these molecules when used at high concentrations.

Despite our failure to detect Hh pathway activity in epithelial cells *in vitro*, we determined that subsets of human tumours and cell lines do indeed express Hh ligands. Microarray expression analysis of human tissue specimens revealed that subsets of colorectal, endometrial, ovarian and pancreatic cancers overexpressed Hh ligand mRNA (data not shown). Quantitative polymerase chain reaction with reverse transcription (RT-PCR) profiling of an independent set of human tissue specimens confirmed that the transcript levels of *SHH* and/or Indian hedgehog homologue (*IHH*) ligands were significantly upregulated in subsets of these cancers (Fig. 2a–c). In contrast, we did not observe increased *SHH* or *IHH* expression in a smaller panel of lung cancer specimens (Fig. 2d). Differential expression of SHH and

¹Genentech Inc., 1 DNA Way, South San Francisco, California 94080, USA. ²Curis Inc., 45 Moulton Street, Cambridge, Massachusetts 02138, USA. [†]Present address: Harvard Stem Cell Institute, Harvard University, Biolabs Room 1065, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA.

*These authors contributed equally to this work.

IHH mRNA and protein was recapitulated in cancer cell lines representative of these tissue types (Supplementary Figs 6 and 7). To determine whether the ligands produced by these cancer cell lines are functional, we co-cultured these cells with 10T1/2 cells stably transfected with a *GLI* reporter construct (Fig. 2e). The induction of *GLI* reporter activity in the fibroblasts correlated with the levels of ligand produced by the tumour cells, demonstrating their capacity to stimulate paracrine signalling. To establish whether a similar paracrine-mediated activation could occur *in vivo*, xenografts were established in *Ptch1-lacZ;Rag2^{-/-}* mice from the pancreatic cell lines HPAF-II and PL45, which differ in expression of Hh ligands (Supplementary Fig. 6). Anti- β -galactosidase staining of Hh-positive HPAF-II xenografts resulted in labelling of stromal cells immediately adjacent to the tumour cell compartment (Fig. 2f and Supplementary Fig. 8). In contrast, β -galactosidase staining was not observed in the stroma of xenografts derived from PL45 cells, which express very low levels of Hh ligands. Taken together, these results demonstrate that

Hh-expressing tumour epithelial cells are capable of activating the Hh pathway in a paracrine manner, both *in vitro* and *in vivo*.

To uncover the role of paracrine Hh signalling in regulation of tumour growth and to ensure that the lack of Hh signalling in cell lines was not a feature of cell culture¹⁵, we first used models in which tumour cells had never been grown *in vitro*. We generated mouse xenograft models from surgical biopsies of patients harbouring a variety of neoplasms. On implantation of these primary tumours

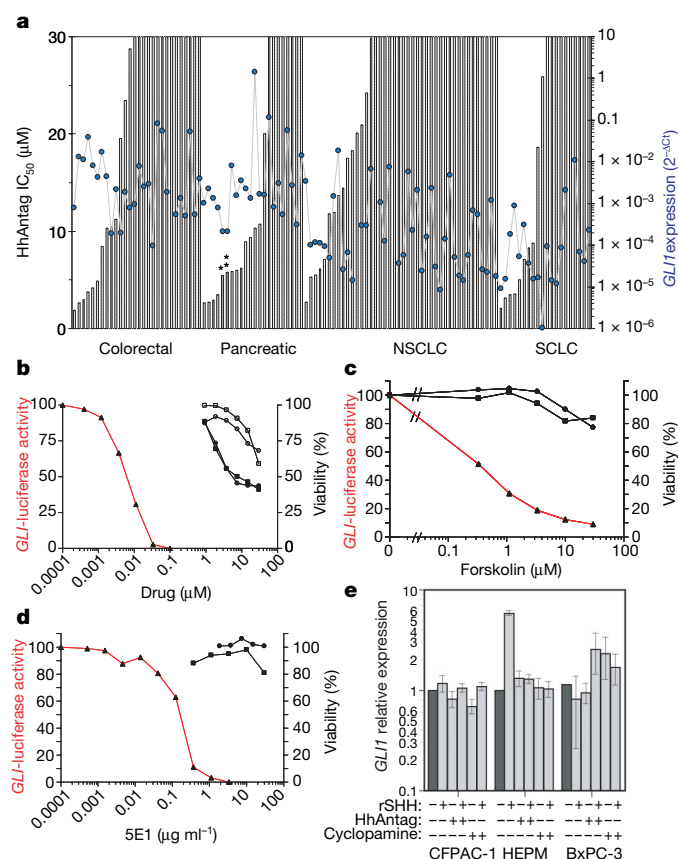


Figure 1 | Lack of Hh pathway activation in human cancer cell lines. **a**, No association was shown between Hh antagonist-mediated growth inhibition (IC₅₀, left y axis, solid bars) and quantitatively determined Hh target gene expression (*GLI1*, right y axis, blue dots). BxPC-3 and CFPAC-1 cells are highlighted by single and double asterisks, respectively. NSCLC, non-small-cell lung cancer; SCLC, small-cell lung cancer. **b–d**, Equivalent degrees of growth inhibition mediated by Hh pathway antagonists in a *SMO*-positive (CFPAC-1, squares) compared to a *SMO*-negative (BxPC-3, circles) cell line. The concentration of antagonists required to inhibit Hh pathway activity was determined by measuring luciferase activity (red line) in a cell line expressing a Hh-responsive *GLI*-luciferase reporter (HEPM-rep, triangles). Hh pathway antagonists include cyclopamine (**b**, open symbols), HhAntag (BxPC-3, filled symbols), forskolin (**c**) and the anti-Hh-ligand-blocking antibody 5E1 (**d**). **e**, Both SHH (rSHH, 1 μg ml⁻¹) and Hh antagonists (10 μM) fail to modulate endogenous Hh target genes in cancer cells. Human embryonic palatal mesenchyme (HEPM) serves as a positive control. Error bars, mean ± s.d. (*n* = 3).

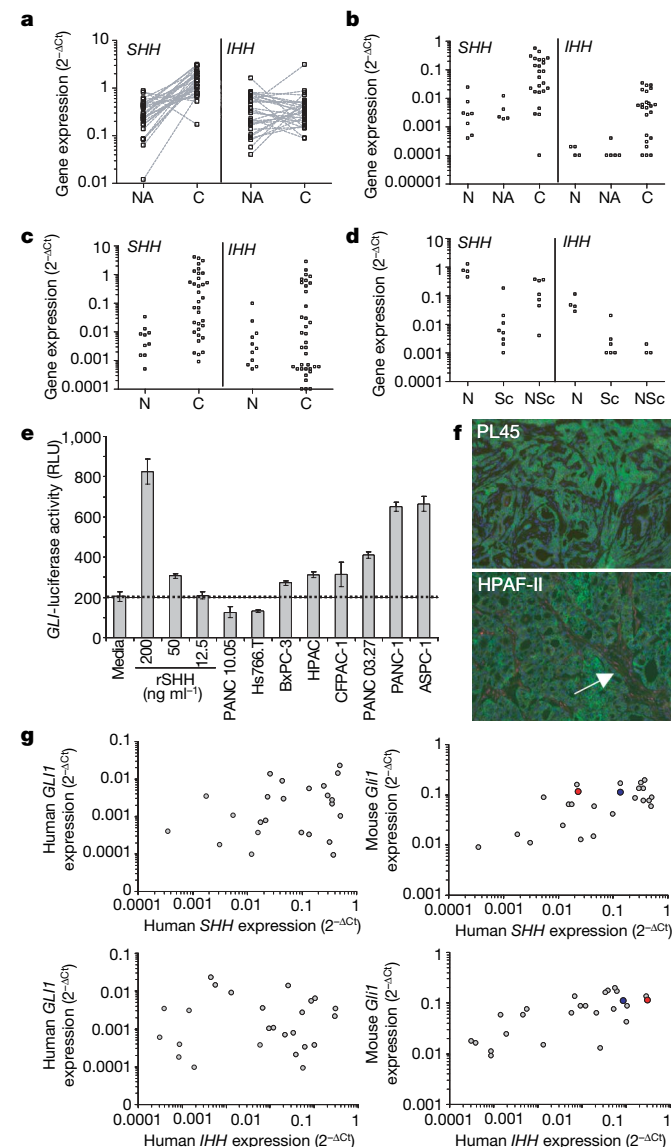


Figure 2 | Human Hh ligands are overexpressed in a subset of human cancers and can activate Hh signalling in a paracrine manner. Quantitative RT-PCR profiling of *SHH* and *IHH* mRNA in a subset of human tissues: **a**, colorectal; **b**, pancreatic; **c**, ovarian; and **d**, lung, SCLC (Sc) and NSCLC (NSc). Colorectal tissue is patient-matched. C, primary tumour tissue; N, normal; NA, normal adjacent. **e**, Induction of Hh-responsive reporter activity in 10T1/2-rep fibroblasts after co-culture with pancreatic cells or rSHH. Error bars, mean ± s.d. (*n* = 3). RLU, relative luciferase units. **f**, Stromal activation of the Hh pathway activity in HPAF-II, but not PL45, xenografts *in vivo*. β -galactosidase activity (red) was observed in the stromal compartment after implantation of HPAF-II cells in *Ptch1-lacZ;Rag2^{-/-}* mice; epithelial-specific antigen is shown in green. Arrow shows β -galactosidase-positive cells. **g**, Correlation between stromal-derived, but not tumour-derived, *GLI1* mRNA levels versus tumour-derived Hh ligand (*SHH* and *IHH*) levels in primary human tumour xenografts (*n* = 22, *P* = 0.005 *SHH/GLI1*; *P* = 0.0003 *IHH/GLI1*) by species-specific qRT-PCR. The red and blue dots refer to the xenograft models in Fig. 3a (blue, D5123; red 1051178-A).

in nude mice, human-derived stroma is rapidly replaced by host mouse stroma (Supplementary Fig. 9)¹⁶. To differentiate Hh pathway activity in the tumour versus stromal compartments, species-specific primer/probe sets were generated to detect human tumour versus mouse stromal pathway genes (Supplementary Table 2). We observed a statistically significant correlation between the levels of *SHH* and *IHH* expression in the tumour and activation of Hh signalling, as judged by *Gli1* (Fig. 2g, right panels) and *Ptch1* (data not shown), in the stroma. This correlation was not detected within the tumour compartment alone (Fig. 2g, left panels, *GLI1*). Human *PTCH1* mRNA was also evaluated by qRT-PCR and no correlation was observed with human *SHH* or *IHH* in the tumour compartment (data not shown). Overall, the absolute level of *Gli1* expression in the stromal infiltrate of individual xenografts was higher compared to those originating from the tumour epithelium (Supplementary Fig. 10). Taken together, these data support a model in which Hh ligands produced by the tumour cells activate the pathway in the surrounding stroma.

To address whether tumours producing Hh ligands are dependent on the paracrine activation of the Hh pathway for growth, xenografts were generated from biopsies of patients with pancreatic adenocarcinoma (1051178-A) and colon adenocarcinoma (D5123). Both models express significant levels of *SHH* and/or *IHH* mRNA (Fig. 2g). D5123 tumours did not express any detectable *SMO* mRNA in the tumour epithelium (cycling threshold, $C_t = 39$). Oral administration of HhAntag to mice harbouring these primary human xenografts resulted in significant growth delay in both models (Fig. 3a), with average tumour growth inhibitions of 29% and 48%, respectively. Significantly, the doses of HhAntag required to inhibit growth were similar to the doses required to fully inhibit endogenous Hh target genes in tumour stroma or in surrogate normal tissues, suggesting that growth inhibition is a specific consequence of Hh inhibition (Supplementary Fig. 11). Next we analysed a panel of colon cell lines with differential levels of Hh ligand expression. Whereas HT55 and HT-29 cells expressed increased levels of both *SHH* and *IHH*, neither ligand was detectable in DLD-1 cells (Supplementary Figs 6 and 7). Furthermore, HT-29 cells did not express detectable *SMO* mRNA ($C_t = 40$). HhAntag treatment resulted in a statistically significant growth delay of HT55 and HT-29 xenografts compared to vehicle controls, whereas HhAntag had no effect on the growth of DLD-1 xenografts (Fig. 3b). Furthermore, the *in vivo* activity was not reflective of the activity observed *in vitro* (Fig. 1a and Supplementary Fig. 5). To demonstrate further that the growth delay observed with HhAntag is specific, mice harbouring HT55 and HT-29 xenografts were treated with the Hh ligand-blocking antibody 5E1. Analogous to the degree of growth inhibition observed with HhAntag, 5E1 resulted in a statistically significant growth retardation in both models (Fig. 3c). In all xenografts tested, Hh pathway antagonist treatment resulted in the downregulation of canonical Hh target genes in the stromal microenvironment, but not within the tumour epithelium, as determined by species-specific expression profiling (Fig. 4a, b and data not shown). Overall, these data suggest that the growth effect mediated by Hh pathway inhibition is by means of its action on the stromal microenvironment and is consistent with a paracrine signalling mechanism.

Finally, to validate further the contribution of Hh-activated stroma to tumour growth, we assessed the ability of fibroblasts deficient in Hh signalling to support the growth of HT-29 cells expressing luciferase (HT29-luc). Mouse embryonic fibroblasts (MEFs) were isolated from mice engineered to conditionally remove *Smo* on tamoxifen induction of Cre recombinase activity (*CAGGCreER*; *Smo*^{C/C})^{17,18}. After tamoxifen treatment, induction of Cre activity led to a greater than 85% reduction in *Smo* expression (Supplementary Fig. 12), and near complete inhibition of *Gli1* upregulation in response to Hh stimulation when compared to control (*Smo*^{C/C}) MEFs (Fig. 4c). Whereas implantation of a suboptimal number of HT29-luc cells in mice did not result in measurable tumours at 26 days

post-implantation, co-injection of HT29-luc cells with tamoxifen-treated MEFs from *CAGGCreER*; *Smo*^{C/C} mice resulted in smaller tumours compared to HT29-luc cells co-injected with *Smo*^{C/C} control MEFs (Fig. 4d). Imaging of mice revealed a significant decrease in luciferase activity in mice implanted with HT29-luc/*CAGGCreER*; *Smo*^{C/C} MEFs compared to HT29-luc/*Smo*^{C/C} MEFs, demonstrating that the tumour size differences were indeed caused by differences in HT29-luc growth (Fig. 4e, f). These data demonstrate that Hh activity in the stromal microenvironment can provide a growth advantage to tumour cells.

We have found that tumours overexpressing Hh ligand(s) activate the signalling pathway in neighbouring stromal cells. This conclusion differs from that presented in several recent publications, the conclusions of which were largely based on inhibition of proliferation using high concentrations of Hh pathway antagonists, as well as the use of artificial reporter constructs to monitor Hh pathway activity. Although we cannot rule out a potential role for Hh signalling in a

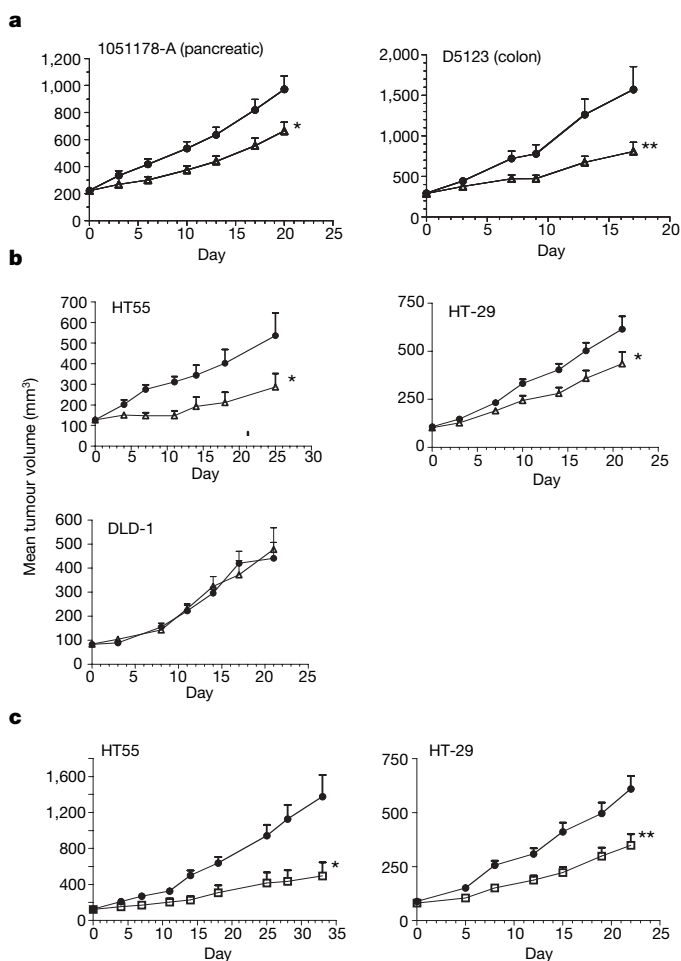


Figure 3 | Inhibition of the Hh pathway results in a significant delay in the growth of Hh ligand-expressing tumours *in vivo*. **a**, Change in tumour volume in two primary human tumour xenograft models after treatment with vehicle (filled circles) or Hh antagonists (open triangles; 75 mg kg⁻¹ twice daily for 1051178-A and 100 mg kg⁻¹ twice daily for D5123); * $P < 0.001$, ** $P < 0.05$ ($n = 11$ per arm, 1051178-A; $n = 8$ per arm, D5123). **b**, Change in tumour volume in three colorectal cell line xenografts after treatment with vehicle (filled circles) or HhAntag (open triangles; 75 mg kg⁻¹ twice daily for HT-29, 100 mg kg⁻¹ twice daily for HT55 and DLD-1); * $P < 0.05$ ($n = 11$ per arm, HT55; $n = 10$ per arm, DLD-1; $n = 20$ vehicle arm, $n = 15$ HhAntag arm, HT-29). **c**, Change in tumour volume in two Hh-positive colorectal cell line xenografts after treatment with an isotype control IgG (filled circles) or 5E1 (open squares); * $P < 0.01$; ** $P < 0.05$ ($n = 10$ per arm, both models). Data in all panels are expressed as mean tumour volumes \pm s.e.m.

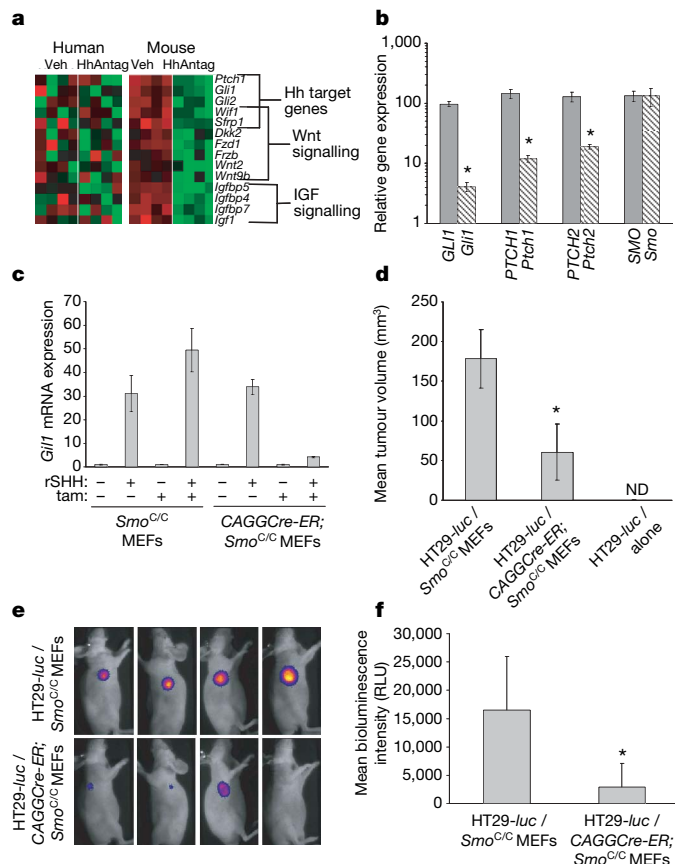


Figure 4 | Stromal hedgehog signalling can support tumour growth. **a**, Heat map representing mRNA expression of selected human and mouse genes in p1051178-A xenografts after treatment with vehicle (Veh) or Hh antagonist (HhAntag). Red, high expression; green, low expression. IGF, insulin-like growth factor. **b**, Hh antagonist treatment results in repression of stromal-derived (hatched bars, * $P < 0.01$), but not tumour-derived (solid bars), Hh target genes in 1051178-A xenografts ($n = 5$), as assessed by species-specific qRT-PCR. **c**, Inability of rSHH stimulation to induce *Gli1* mRNA in tamoxifen-treated MEFs derived from CAGGCre-ER; *Smo*^{C/C} mice compared to *Smo*^{C/C} MEFs ($n = 3$). **d**, Reduction in tumour size 26 days after the implantation of 3×10^5 HT29-luc cells with 1.5×10^6 of tamoxifen-treated CAGGCre-ER; *Smo*^{C/C} MEFs compared to *Smo*^{C/C} MEFs. ND, not detected ($n = 4$). **e**, In vivo imaging of HT29-luc tumour cells in **d**. **f**, Quantified bioluminescent images of HT29-luc tumour cells in **e** (* $P = 0.04$). All data are expressed as means \pm s.d.

minor subpopulation of epithelial tumour cells such as self-renewing cancer stem cells¹⁹, an epithelia-to-mesenchyma signalling paradigm has been well described for Hh in various non-neoplastic settings⁹ and data are beginning to emerge that support this paradigm in neoplastic models²⁰. How Hh pathway inhibition in the stromal microenvironment suppresses tumour growth remains to be elucidated. Hh-driven mesenchymal signals could influence the structural interactions between the tumour epithelium and its microenvironment or stromal Hh signalling could affect the expression of soluble factors in the tumour microenvironment that are used by either the tumour epithelium or additional stromal cell types. Stimulation of the Hh pathway in mesenchymal cells has been shown to regulate genes and pathways that are involved in tumorigenesis and metastasis^{21,22}. We find that inhibition of Hh signalling in the stroma can modulate factors, such as insulin-like growth factor receptor and Wnt signalling pathway components (Fig. 4a) that have been linked to the pathogenesis of these respective tumour types^{23–26}; however, the exact role of these pathway(s) remains to be determined. Given the growing appreciation of the role that the tumour microenvironment has in regulating the progression of cancer^{27,28}, it is probable that a combination of multiple factors modulated by Hh pathway

inhibition in the stroma would affect the growth characteristics of these tumours. Our results have important implications for the clinical development of Hh pathway antagonists and present a new understanding of the way in which Hh antagonists may slow tumour growth.

METHODS SUMMARY

Cell culture assays. Cells were cultured in triplicate in 0.5% serum-containing medium containing HhAntag (Supplementary Fig. 13), cyclopamine or the anti-SHH monoclonal antibody 5E1 (ref. 29). The concentration of antagonists resulting in 50% inhibition of cell viability at 72 h was determined. Co-culture experiments were carried out by culturing tumour cells with C3H10T1/2 fibroblasts expressing luciferase under a *Gli1*-responsive promoter for 24 h. Assessment of endogenous Hh pathway gene expression was performed by culturing cells for 24 h with rSHH and/or Hh antagonists in 0.5% serum-containing medium.

In vivo models. Primary human xenografts were established by direct implantation of surgical material into female CD1 nu/nu mice of 6–8 weeks of age and were serially passaged into larger cohorts of mice for efficacy testing on reaching 200–350 mm³. HT-29, HT55 and DLD-1 cell lines were established as xenografts by injection into the hind flank of CD1 nu/nu mice with 5×10^6 (HT-29 and HT55) or 10×10^6 (DLD-1) cells. HhAntag, or a close derivative, was resuspended in 0.5% methyl-cellulose, 0.2% Tween-80 (MCT) and administered orally twice daily as indicated. 5E1 or an isotype control (IgG1) were administered at 60 mg kg⁻¹ on day 1 followed by 30 mg kg⁻¹ weekly by means of intraperitoneal delivery. Tumour volumes were calculated as $(L \times W \times W)/2$, where W represents width and L represents length. Statistical analysis was performed on mean tumour volumes at the end of the study using the Dunnett's test. Conditional (floxed) *Smo* mice (*Smo*^{C/C}) and tamoxifen-inducible CAGGCre-ER transgenic mice were obtained from A. McMahon.

Gene expression profiling. Microarray gene expression analysis was carried out on Affymetrix Human Genome U133 Plus 2.0 (Supplementary Information File 1) or Mouse Genome 430 2.0 arrays (Supplementary Information File 2). Hh pathway genes were quantitatively assessed in frozen human tissue lysates, cell lines or xenograft tissues by Taqman. Assay information can be found in Supplementary Tables 2 and 3.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 30 April; accepted 14 July 2008.

Published online 27 August 2008.

- Clement, V., Sanchez, P., de Tribolet, N., Radovanovic, I. & Ruiz i Altaba, A. HEDGEHOG-GLI1 signaling regulates human glioma growth, cancer stem cell self-renewal, and tumorigenicity. *Curr. Biol.* **17**, 165–172 (2007).
- Stecca, B. et al. Melanomas require HEDGEHOG-GLI1 signaling regulated by interactions between GLI1 and the RAS-MEK/AKT pathways. *Proc. Natl Acad. Sci. USA* **104**, 5895–5900 (2007).
- Yuan, Z. et al. Frequent requirement of hedgehog signaling in non-small cell lung carcinoma. *Oncogene* **26**, 1046–1055 (2007).
- Thayer, S. P. et al. Hedgehog is an early and late mediator of pancreatic cancer tumorigenesis. *Nature* **425**, 851–856 (2003).
- Karhadkar, S. S. et al. Hedgehog signalling in prostate regeneration, neoplasia and metastasis. *Nature* **431**, 707–712 (2004).
- Berman, D. M. et al. Widespread requirement for Hedgehog ligand stimulation in growth of digestive tract tumours. *Nature* **425**, 846–851 (2003).
- Watkins, D. N. et al. Hedgehog signalling within airway epithelial progenitors and in small-cell lung cancer. *Nature* **422**, 313–317 (2003).
- Ingham, P. W. & McMahon, A. P. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* **15**, 3059–3087 (2001).
- van den Brink, G. R. Hedgehog signaling in development and homeostasis of the gastrointestinal tract. *Physiol. Rev.* **87**, 1343–1375 (2007).
- Romer, J. T. et al. Suppression of the Shh pathway using a small molecule inhibitor eliminates medulloblastoma in *Ptc1*^{+/+} *p53*^{-/-} mice. *Cancer Cell* **6**, 229–240 (2004).
- Frank-Kamenetsky, M. et al. Small-molecule modulators of Hedgehog signaling: identification and characterization of Smoothened agonists and antagonists. *J. Biol.* **1**, 10 (2002).
- Cooper, M. K., Porter, J. A., Young, K. E. & Beachy, P. A. Teratogen-mediated inhibition of target tissue response to Shh signaling. *Science* **280**, 1603–1607 (1998).
- Kayed, H. et al. Indian hedgehog signaling pathway: expression and regulation in pancreatic cancer. *Int. J. Cancer* **110**, 668–676 (2004).
- Fan, C. M. et al. Long-range sclerotome induction by sonic hedgehog: direct role of the amino-terminal cleavage product and modulation by the cyclic AMP signaling pathway. *Cell* **81**, 457–465 (1995).

15. Sasai, K. *et al.* Shh pathway activity is down-regulated in cultured medulloblastoma cells: implications for preclinical studies. *Cancer Res.* **66**, 4215–4222 (2006).
16. Duda, D. G. *et al.* Differential transplantability of tumor-associated stromal cells. *Cancer Res.* **64**, 5920–5924 (2004).
17. Hayashi, S. & McMahon, A. P. Efficient recombination in diverse tissues by a tamoxifen-inducible form of Cre: a tool for temporally regulated gene activation/inactivation in the mouse. *Dev. Biol.* **244**, 305–318 (2002).
18. Long, F., Zhang, X. M., Karp, S., Yang, Y. & McMahon, A. P. Genetic manipulation of hedgehog signaling in the endochondral skeleton reveals a direct role in the regulation of chondrocyte proliferation. *Development* **128**, 5099–5108 (2001).
19. Beachy, P. A., Karhadkar, S. S. & Berman, D. M. Tissue repair and stem cell renewal in carcinogenesis. *Nature* **432**, 324–331 (2004).
20. Fan, L. *et al.* Hedgehog signaling promotes prostate xenograft tumor growth. *Endocrinology* **145**, 3961–3970 (2004).
21. Pola, R. *et al.* The morphogen Sonic hedgehog is an indirect angiogenic agent upregulating two families of angiogenic growth factors. *Nature Med.* **7**, 706–711 (2001).
22. Ingram, W. J., Wicking, C. A., Grimmond, S. M., Forrest, A. R. & Wainwright, B. J. Novel genes regulated by Sonic Hedgehog in pluripotent mesenchymal cells. *Oncogene* **21**, 8196–8205 (2002).
23. Bergmann, U., Funatomi, H., Yokoyama, M., Beger, H. G. & Korc, M. Insulin-like growth factor I overexpression in human pancreatic cancer: evidence for autocrine and paracrine roles. *Cancer Res.* **55**, 2007–2011 (1995).
24. Durai, R., Yang, W., Gupta, S., Seifalian, A. M. & Winslet, M. C. The role of the insulin-like growth factor system in colorectal cancer: review of current knowledge. *Int. J. Colorectal Dis.* **20**, 203–220 (2005).
25. Zeng, G. *et al.* Aberrant Wnt/ β -catenin signaling in pancreatic adenocarcinoma. *Neoplasia* **8**, 279–289 (2006).
26. Pasca di Magliano, M. *et al.* Common activation of canonical wnt signaling in pancreatic adenocarcinoma. *PLoS ONE* **2**, e1155 (2007).
27. Bhowmick, N. A., Neilson, E. G. & Moses, H. L. Stromal fibroblasts in cancer initiation and progression. *Nature* **432**, 332–337 (2004).
28. Orimo, A. & Weinberg, R. A. Stromal fibroblasts in cancer: a novel tumor-promoting cell type. *Cell Cycle* **5**, 1597–1601 (2006).
29. Ericson, J., Morton, S., Kawakami, A., Roelink, H. & Jessell, T. M. Two critical periods of Sonic Hedgehog signaling required for the specification of motor neuron identity. *Cell* **87**, 661–673 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank A. McMahon and M. Scott for providing transgenic mice; T. Holcomb, K. Wagner, D. Lee and P. Wen for their assistance in cell line screening; J. Ernst for rSHH, M. Cole for assistance with imaging; S. Louie for assistance with graphics; P. Haverly for assistance with gene expression data; and M. Evangelista, C. Callahan and V. Dixit for comments and discussions. Tissue samples were provided by the Cooperative Human Tissue Network, which is funded by the National Cancer Institute, and the National Center for Research Resources, which is supported by the National Institutes of Health. Other investigators may have received samples from these same tissues.

Author Contributions F.J.d.S. and L.L.R. conceived and directed the project. S.E.G., R.L.Y., S.J.S., T.T., H.T. J.C.M. and K.K. designed and carried out experiments. C.P.A., D.M., L.F., T.J., D.K. and M.N.-P. carried out experiments. R.L.Y., S.E.G. and F.J.d.S. wrote the paper.

Author Information Microarray data are deposited in the NCBI GEO database under accession number GSE11981. The authors declare competing financial interests: details accompany the full-text HTML version of the paper at www.nature.com/nature. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.J.d.S. (sauvage@gene.com).

METHODS

Reagents, cell lines and cell culture. HhAntag was synthesized as described in Supplementary Fig. 13. Cycloplamine was purchased from Toronto Research Chemicals, Inc. The mouse anti-Shh monoclonal antibody 5E1 (ref. 29) was protein-A-purified and confirmed for lack of aggregation and endotoxin, as well as the ability to block ligand-activated Hh signalling, after gel filtration.

HEPM cells³⁰ were stably transfected with a luciferase reporter gene driven by six continuous repeats of the *GLI* DNA-binding element. The human colon fibroblast line CCD-18Co was purchased from ATCC (CRL-1459). MEFs from *CAGGCre-ER;Smo^{C/C}* and control mice were isolated according to published methods and treated with tamoxifen (1 μ M) *in vitro* for 5 days before testing for Hh responsiveness *in vitro* and/or mixing with HT29-*luc* cells for implantation into CD1 nude mice.

Cell viability was measured at 72 h using the Celltiter-Glo Luminescent Cell Viability Assay kit (Promega), and the concentration of HhAntag resulting in 50% inhibition of cell viability was determined from a minimum of two experiments. Co-culture experiments were carried out by culturing tumour cells at a 4:1 ratio with C3H10T1/2 S12 fibroblast *GLI*-reporter cells¹¹. Luciferase activity was measured using the SteadyLite HTS kit (Promega) after 24 h. Assessment of Hh pathway target gene expression was performed by culturing cells in triplicate in 24-well plates for 24 h with 300 ng ml⁻¹ rSHH and/or Hh antagonists in 0.5% serum-containing medium. RNA was isolated using Qiagen RNeasy Mini Kit.

Reporter gene assays. Twenty-four hours after plating, cells were transfected for 18 h with either *GLI*-luciferase or *NF κ B*-luciferase reporter plasmids in combination with a control herpes simplex virus thymidine kinase (*HSV-TK*)-luciferase plasmid using the Eugene6 transfection reagent (Roche) and then re-plated into 96-well culture plates. Six hours later, Hh antagonist and/or 1 μ g ml⁻¹ rSHH were added to cells in quadruplicate in 0.5% serum-containing medium. Cultures were re-fed 24 h later, and plates were assessed for firefly and *Renilla* luciferase activity after an additional 24 h using the Promega Dual-Glo luciferase kit.

Immunofluorescence. To evaluate Hh pathway activation *in vivo*, *Ptch1-lacZ;Rag2^{-/-}* mice were implanted with pancreatic tumour cell lines expressing differential levels of Hh ligands. Fourteen days after implantation, xenografts were excised and fixed in 4% paraformaldehyde before embedding in OCT (frozen tissue matrix). Sections were incubated with anti- β -galactosidase (1:10,000, Cappel) and FITC-conjugated anti-ESA (epithelial specific antigen, 1:100, Biomed) overnight (16 h), followed by secondary Cy3-anti-rabbit (1:400, Jackson ImmunoResearch) incubation for one hour. Nuclei were visualized with 4,6-diamidino-2-phenylindole (DAPI, blue) in Vectamount (Vector Laboratory).

Xenograft models. Primary tumour samples were provided by the National Disease Research Interchange (NDRI) and the Cooperative Human Tissue Network, which is funded by the National Cancer Institute. Other investigators may have received samples from these same tissues. Primary human xenografts were established by direct implantation of surgical material into female CD1 nu/nu mice of 6–8 weeks of age (Charles River Laboratories, Inc.). All mice were housed and maintained according to the animal use guidelines of Genentech,

Inc., conforming to California State legal and ethical practices. Tissue specimens were shipped in RPMI-containing antibiotics on wet ice and implanted within 24 h after rinsing in DMEM containing 0.11 mg ml⁻¹ sodium pyruvate, 1.125 μ g ml⁻¹ amphotericin B and 1 mg ml⁻¹ kanamycin. Tissue was minced with scalpels to a size of <1 mm³, and approximately 100 mg of tissue was implanted in the subcutaneous space of the hind flank using blunt dissection and a 10G trocar. Tumour lines were serially passaged into larger cohorts of mice for efficacy testing. Mice were distributed into tumour-volume-matched cohorts when tumours reached between 200 mm³ and 350 mm³. HT-29, HT55 and DLD-1 cell lines were purchased from ATCC and were established as xenografts by injection into the hind flank of 6–8-week-old female CD1 nu/nu mice with 5×10^6 (HT-29 and HT55) or 10×10^6 (DLD-1) cells resuspended in Hanks' balanced salt solution. Tumour-bearing mice were distributed into tumour-volume-matched cohorts when the tumours reached between 80 mm³ and 120 mm³. HhAntag, or a close derivative, was resuspended in MCT and administered orally twice daily at either 100 or 75 mg kg⁻¹ from a 10 mg ml⁻¹ suspension as indicated. 5E1 or an isotypic control (IgG1) were dissolved in PBS and administered at 60 mg kg⁻¹ on day 1 and then at 30 mg kg⁻¹ weekly by intraperitoneal delivery. Tumour volume and animal weights were monitored twice weekly and tumour volume calculated as (L \times W \times W)/2.

Expression studies. For the analysis of *SHH* and *IHH* mRNA expression in multiple human tissue specimens, data were obtained from Gene Logic, Inc.

Microarray gene expression analysis of RNA extracted from primary xenograft tissue was carried out on two separate platforms, Human Genome U133 Plus 2.0 and Mouse Genome 430 2.0 arrays (Affymetrix). Preparation of complementary RNA, array hybridizations and subsequent data analysis were carried out using the manufacturers' protocols, with signal intensities being determined by the MAS5.0 algorithm. Hh pathway genes were quantitatively assessed by Taqman and transcript levels were normalized to the housekeeping genes β -glucuronidase (*GUSB*) or ribosomal protein L19 (*RPL19*). Results are expressed as normalized expression values ($= 2^{-\Delta\Delta Ct}$) or normalized expression relative to a cell line pool representative of multiple tissue types ($= 2^{-\Delta\Delta Ct}$), unless otherwise stated. For xenograft model profiling, gene expression in each compartment was normalized to the same species-specific housekeeping gene to ensure that observed differences weren't due to differences in the degree of stromal infiltrate. Correlations in gene expression were evaluated by Spearman rank tests and *P*-values were reported. Sequences of primer/probes are shown in Supplementary Table 3 and species-specificity of human/mouse Taqman assays is confirmed in Supplementary Table 2.

Fluorescence-activated cell sorting. Cells were stained with either mouse anti-Hh antibody 5E1 or isotype control antibody and then followed by biotin-conjugated anti-mouse, streptavidin-phycoerythrin and propidium iodide. Stained live (PI-excluded) cells were analysed on a FACS Calibur, and data were plotted using the FlowJo software package.

30. Yoneda, T. & Pratt, R. M. Mesenchymal cells from the human embryonic palate are highly responsive to epidermal growth factor. *Science* **213**, 563–565 (1981).

S-nitrosylation of histone deacetylase 2 induces chromatin remodelling in neurons

Alexi Nott¹, P. Marc Watson¹, James D. Robinson¹, Luca Crepaldi¹ & Antonella Riccio¹

Brain-derived neurotrophic factor (BDNF) and other neurotrophins have a vital role in the development of the rat and mouse nervous system by influencing the expression of many specific genes that promote differentiation, cell survival, synapse formation and, later, synaptic plasticity¹. Although nitric oxide (NO) is known to be an important mediator of BDNF signalling in neurons², the mechanisms by which neurotrophins influence gene expression during development and plasticity remain largely unknown. Here we show that BDNF triggers NO synthesis and S-nitrosylation of histone deacetylase 2 (HDAC2) in neurons, resulting in changes to histone modifications and gene activation. S-nitrosylation of HDAC2 occurs at Cys 262 and Cys 274 and does not affect deacetylase activity. In contrast, nitrosylation of HDAC2 induces its release from chromatin, which increases acetylation of histones surrounding neurotrophin-dependent gene promoters and promotes transcription. Notably, nitrosylation of HDAC2 in embryonic cortical neurons regulates dendritic growth and branching, possibly by the activation of CREB (cyclic-AMP-responsive-element-binding protein)-dependent genes. Thus, by stimulating NO production and S-nitrosylation of HDAC2, neurotrophic factors promote chromatin remodelling and the activation of genes that are associated with neuronal development.

Although most cells in multicellular organisms contain an identical genome, different cell types stably express different sets of genes, reflecting progressive and persistent alterations in chromatin structure that underlie cell specialization and cell memory^{3,4}. The differentiation of multipotent stem cells into neurons is a quintessential example of epigenetics at work. Despite growing evidence that neurotrophic factors and other extracellular signals influence the epigenetic changes associated with gene transcription, the details of how they regulate these changes are still poorly understood. Neurotrophins are a family of peptide growth factors that exert their growth and survival-promoting effects on neurons through the activation of various intracellular signalling pathways, including the NO pathway^{2,5,6}. In the nervous system, NO has been linked to many physiological and pathological functions, including neurogenesis, hippocampal long-term potentiation, neurodegeneration and neuronal survival and differentiation^{7,8}. There is increasing evidence that NO can act by directly modifying cysteine residues in target proteins^{9,10}. Cysteine nitrosylation (S-nitrosylation) is now regarded as a selective and specific signal controlled by cellular NO⁹. S-nitrosylation of transcription factors in the cytoplasm, for example, is one way that NO can regulate gene expression¹¹.

To investigate how neurotrophins regulate NO-dependent gene expression, we first tested whether BDNF stimulation induced nuclear localization of NO in rat cortical neurons. To visualize NO, we used 4-amino-5-methylamino-2',7'-difluororescein diacetate (DAF-FM DA), a fluorescent dye that specifically binds NO¹². In basal conditions, NO was synthesized at low levels in many neurons

(Fig. 1a, b). Stimulation of cortical neurons with BDNF for 30 min resulted in a rapid accumulation of DAF-FM DA reactivity in both cytoplasmic and nuclear compartments. Pre-treatment of cortical neurons with the specific neuronal NO synthase (nNOS) inhibitor NO-propyl-L-arginine (NPA, 300 μ M) consistently resulted in a lack of BDNF-stimulated DAF-FM DA reactivity (Fig. 1a, b). Notably, removal of NPA from the culture medium followed by stimulation with BDNF resulted in a pronounced increase of DAF-FM DA reactivity in both the nuclear and cytoplasmic compartments (Fig. 1a, b). To test whether the NO generated on neurotrophin treatment directly modifies cytoplasmic and nuclear proteins by means of S-nitrosylation, we used the biotin-switch assay^{12–17}. Stimulation of cortical neurons with BDNF induced S-nitrosylation of several proteins detected in whole cell extracts (Fig. 1c, arrows), many of which were nitrosylated within minutes of exposure of the neurons to BDNF and remained nitrosylated 1 h after stimulation. Measurement of nuclear and cytoplasmic extracts exposed to the NO donor S-nitrosocysteine (SNOC) showed that several nuclear proteins are targets of NO, and are therefore potential substrates of S-nitrosylation *in vivo* (Fig. 1d, arrows and Supplementary Information 1). The S-nitrosylation pattern of nuclear proteins was clearly distinct from the cytoplasmic pattern, with prominent nuclear nitrosylation of proteins of molecular masses 90, 75, 38 and 25 kDa.

One function of NO in the nucleus may be to regulate the interaction of transcription factors with chromatin^{2,18}. During transcriptional activation, the tightly compacted chromatin is modified by histone acetyltransferases, which are recruited to gene promoters, and this chemical modification is counteracted by histone deacetylases (HDACs)^{19,20}. As NO influences gene expression primarily by changing the acetylation state of chromatin associated with the promoter of neurotrophin-regulated genes² (Supplementary Fig. 2a, b), we investigated whether NO influences histone acetylation by modifying histone deacetylase activity. Messenger RNA and protein expression analyses showed that HDAC2 is highly expressed in embryonic neurons (A.N. and A.R., unpublished observations). To test whether HDAC2 itself may be a target of BDNF-dependent S-nitrosylation, cortical neurons were stimulated with BDNF and nitrosylated proteins were precipitated and subjected to HDAC2 immunoblotting analysis (Fig. 2a). BDNF was found to induce a rapid and sustained S-nitrosylation of HDAC2. Moreover, treatment of either cortical neurons or HEK293T cells overexpressing nNOS with the Ca²⁺ ionophore A23187 also led to an increase in the S-nitrosylation of HDAC2 (Fig. 2a, b). BDNF and A23187 stimulation did not induce HDAC2 S-nitrosylation in cortical neurons lacking nNOS (Fig. 2b and Supplementary Fig. 2d). Experiments performed using the Saville–Griess assay, an alternative biochemical assay to assess protein S-nitrosylation^{21,22}, confirmed that recombinant HDAC2 is strongly nitrosylated after treatment with SNOC *in vitro* (Supplementary Fig. 2c).

¹MRC Laboratory for Molecular and Cell Biology, and Department of Neuroscience, Physiology and Pharmacology, University College London, London WC1E 6BT, UK.

HDAC2 belongs to the class I HDACs, which includes HDAC1 to HDAC3 and HDAC8. Structural analysis has shown that the conserved deacetylase domain of class I HDACs includes an active site containing a zinc-binding domain²³ (Supplementary Fig. 3a). We first tested whether the cysteine located in the deacetylase domain, Cys 152, was the acceptor site of S-nitrosylation. Mutation of Cys 152 to alanine did not influence HDAC2 nitrosylation (A.N. and A.R., unpublished observations). Structure-to-function analysis of

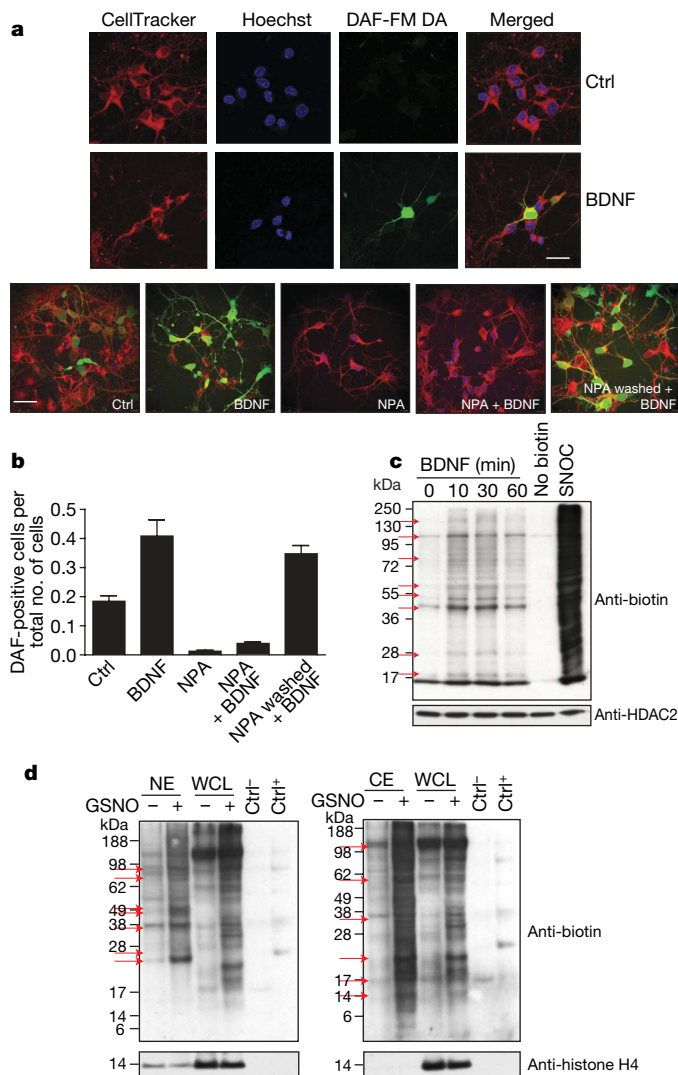


Figure 1 | BDNF induces nuclear NO synthesis and S-nitrosylation of nuclear and cytoplasmic proteins. **a**, Neurons were incubated for 30 min with DAF-FM DA and stimulated with BDNF (75 ng ml⁻¹ for 30 min), NPA (300 μM for 30 min), BDNF in the presence of NPA (NPA + BDNF), NPA followed by BDNF (NPA washed + BDNF) or vehicle (Ctrl). Cells were stained with Hoechst and CellTracker and images were captured using a Leica DM 2500 confocal microscope. Scale bars, 25 μm (top panels) and 75 μm (bottom panel). **b**, Quantitative analysis of the data shown in **a**. Shown are the averages and the s.e.m.; *n* = 3. **c**, Biotin-switch assay of cortical neurons treated with either BDNF for the indicated times (in minutes) or 200 μM SNOC for 15 min. Neurons were subjected to the biotin-switch assay followed by anti-biotin western blot. The arrows indicate proteins that are nitrosylated after BDNF stimulation; *n* = 3. Equal loading was assessed by HDAC2 western blotting. **d**, Neuronal nuclear (left blot) and cytoplasmic (right blot) proteins were exposed to S-nitrosoglutathione (GSNO, 100 μM) *in vitro* and subjected to the biotin-switch assay. The purity of nuclear and cytoplasmic fractions was assessed by western blot analysis of histone H4 (bottom panels). The arrows indicate proteins that are nitrosylated after BDNF stimulation. CE, cytoplasmic extracts; Ctrl⁺, mix of control proteins exposed to GSNO; Ctrl⁻, mix of control proteins left untreated (the mix does not contain histones); NE, nuclear extracts; WCL, whole cell lysate; *n* = 3.

HDAC2 showed that two residues, Cys 262 and Cys 274, were nitrosylated after SNOC treatment (Fig. 2c and Supplementary Fig. 3b, c). Single mutation of either Cys 262 or Cys 274 to alanine did not change HDAC2 nitrosylation, whereas double mutation of both Cys 262 and Cys 274 completely abolished S-nitrosylation of HDAC2 (Fig. 2c). Notably, Cys 262 and Cys 274 were also nitrosylated when cortical neurons were stimulated with BDNF (Fig. 2d, e).

As for other post-translational modifications, S-nitrosylation may regulate protein functions in more than one way. Nitrosylation of critical cysteine residue(s), for example, may influence their enzymatic activity^{16,24,25}. Alternatively, S-nitrosylation may induce a conformational change that influences the interaction of the nitrosylated protein with its binding partners^{12,15,17,26}. To test whether BDNF-dependent NO signalling regulates HDAC2 activity, cortical neurons were stimulated with BDNF or SNOC, cell lysates were immunoprecipitated with an HDAC2 antibody, and HDAC2 activity was assessed using a fluorimetric assay. We observed no difference in HDAC2 activity between untreated and stimulated cells (Supplementary Fig. 4a, b). Moreover, HDAC2 activity was not significantly affected in PC12 cells expressing the wild-type (HDAC2^{WT}) or the double mutant form of HDAC2 (HDAC2^{C262A/C274A}), which is not nitrosylated, after treatment with either nerve growth factor (NGF) or SNOC (Fig. 3a and Supplementary Fig. 4a, b).

To determine whether HDAC2 S-nitrosylation affects its ability to associate with DNA, we performed chromatin immunoprecipitation (ChIP) assays using PC12 cells expressing either HDAC2^{WT} or

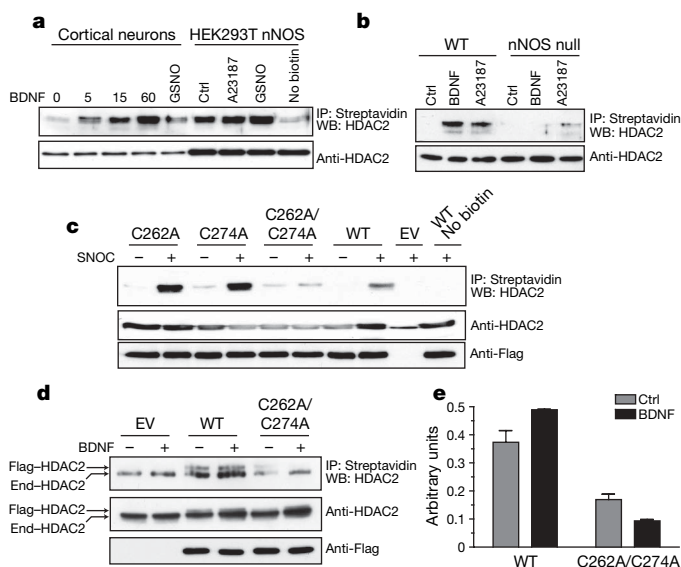


Figure 2 | BDNF induces S-nitrosylation of HDAC2 on Cys 262 and Cys 274. **a**, Cortical neurons and nNOS-expressing HEK293T cells were stimulated with either BDNF for the indicated times (minutes), GSNO or 50 μM A23187 for 20 min. Protein extracts were subjected to the biotin-switch assay followed by streptavidin precipitation and HDAC2 western blotting; *n* = 3. Equal loading was assessed by HDAC2 western blotting. **b**, Biotin-switch assay of BDNF-treated cortical neurons derived from wild-type (left lanes) or nNOS null (right lanes) mouse embryos. Neurons were stimulated with BDNF or A23187 for 20 min. Precipitations were performed using streptavidin followed by HDAC2 western blotting; *n* = 3. **c**, Biotin-switch assay of HEK293T cells transfected with HDAC2^{WT}, HDAC2^{C262A}, HDAC2^{C274A}, HDAC2^{C262A/C274A} or empty vector (EV), and treated with SNOC. Streptavidin precipitation (IP) was followed by HDAC2 western blotting (WB); *n* = 3. **d**, Cortical neurons were transfected with HDAC2^{WT} or HDAC2^{C262A/C274A} vectors and stimulated with BDNF or unstimulated (Ctrl). S-nitrosylation was assessed using the biotin-switch assay. End-HDAC2, endogenous HDAC2; Flag-HDAC2, Flag-tagged HDAC2 expression vectors. **e**, Densitometric analysis of the data in **d**. Shown are the averages and the s.e.m.; *n* = 3.

HDAC2^{C262A/C274A}. We analysed rat *Fos*, *Egr1*, *Vgf* and *Nos1* (also known as *nNOS*) genes, as they are regulated by BDNF and contain a CREB binding site in their promoters². In resting conditions, HDAC2^{WT} was associated with *Fos*, *Egr1*, *Vgf* and *Nos1* promoters. On stimulation with either NGF or SNOC, HDAC2^{WT} readily dissociated from the chromatin (Fig. 3b, c and Supplementary 5a). In contrast, HDAC2^{C262A/C274A} remained strongly associated with the promoters after NGF stimulation or SNOC treatment (Fig. 3b, c and Supplementary 5a). To test whether expression of HDAC2^{C262A/C274A} in PC12 cells inhibited NGF-dependent histone acetylation, we performed ChIP analysis of cells transfected with HDAC2^{WT} or the HDAC2^{C262A/C274A} mutant. As expected, the treatment of PC12 cells expressing either an empty vector or HDAC2^{WT} with NGF or SNOC induced rapid acetylation of histones H3 and H4, as assessed using either acetylated histone H3 K9/K14 or pan-acetylated histone H4 antibodies (Fig. 3d and Supplementary 5b). In contrast, when PC12 cells were transfected with HDAC2^{C262A/C274A}, NGF failed to induce histone acetylation of the

chromatin surrounding the *Fos*, *Egr1*, *Vgf* and *Nos1* promoters (Fig. 3d and Supplementary 5b). Taken together, these results indicate that neurotrophin-dependent nitrosylation of HDAC2 on the critical cysteine residues Cys 262 and Cys 274 is necessary to induce the dissociation of HDAC2 from chromatin and thereby facilitate acetylation of histones H3 and H4.

To determine whether HDAC2 nitrosylation mediates BDNF-dependent transcriptional activation of *Fos*, *Vgf* and *Egr1* genes in cortical neurons, PC12 cells were transfected with HDAC2^{WT}, HDAC2^{C262A/C274A} or an empty vector and stimulated with NGF for various time points. Northern blot analysis showed that the expression of HDAC2^{C262A/C274A} significantly inhibited NGF-dependent transcription of *Fos*, *Vgf* and *Egr1*, whereas HDAC2^{WT} or empty vector did not affect *Vgf* and *Egr1* transcription (Fig. 3e). Although HDAC2^{WT} did not change *Fos* promoter activity (Supplementary Fig. 6a), it did reduce endogenous *Fos* expression, suggesting that it is also affecting chromatin regions outside the CRE-containing promoter.

We next examined the effects of HDAC2 nitrosylation on neuronal differentiation and the development of dendrites. Cortical neurons were transfected with HDAC2^{C262A/C274A}, HDAC2^{WT} or empty vector and the dendritic morphology was visualized by co-transfecting with a green fluorescent protein (GFP) vector. Quantitative analysis showed that in unstimulated neurons, expression of HDAC2^{C262A/C274A} decreased the total and the average dendritic length, whereas dendritic branching was only slightly reduced (Fig. 4a, c and Supplementary 7a). To test whether HDAC2 nitrosylation influenced stimulus-dependent

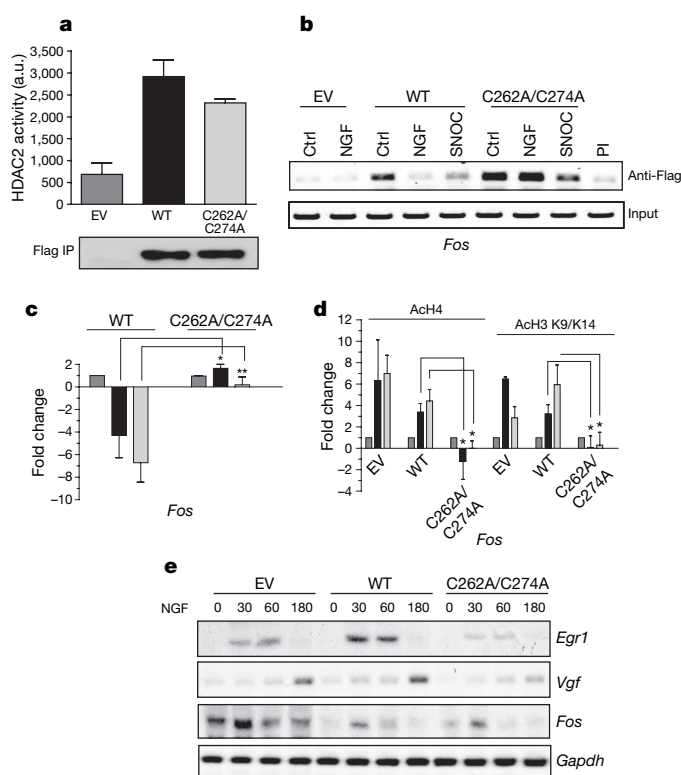


Figure 3 | S-nitrosylation of HDAC2 regulates its association with chromatin. **a**, PC12 cells were transfected with the indicated HDAC2 carboxy-terminal Flag-tagged vectors. Anti-Flag antibody immunoprecipitates (IP) were subjected to HDAC assay. Shown are the averages and the s.e.m.; $n = 3$; a.u., arbitrary units. **b**, ChIP analysis of PC12 cells transfected with Flag-tagged HDAC2^{WT}, HDAC2^{C262A/C274A} or empty vector (EV) and stimulated with NGF (100 ng ml⁻¹) or SNOC for 20 min. Flag immunoprecipitation was followed by PCR for the *Fos* promoter. Ctrl, vehicle; PI, immunoprecipitation with pre-immune serum; $n = 3$. **c**, Real-time PCR quantification of ChIP analyses. Data were normalized by total input and represented as the fold change over unstimulated control. Shown are the averages and the s.e.m.; $n = 3$; asterisk, $P < 0.05$; double asterisk, $P < 0.005$. **d**, Real-time PCR quantification of ChIP analysis of PC12 cells transfected with HDAC2 vectors as indicated, and stimulated with NGF or SNOC. Acetylated histone H3 (AcH3 K9/K14) and pan-acetylated histone H4 (AcH4) immunoprecipitation was followed by quantitative PCR analysis of the *Fos* promoter. Data were normalized by total input and represented as fold changes over unstimulated control. Shown are the averages and the s.e.m.; $n = 3$; asterisk, $P < 0.001$. **e**, Northern blot analysis of PC12 cells transfected with HDAC2 vectors and stimulated with NGF for the indicated times (minutes); $n = 3$.

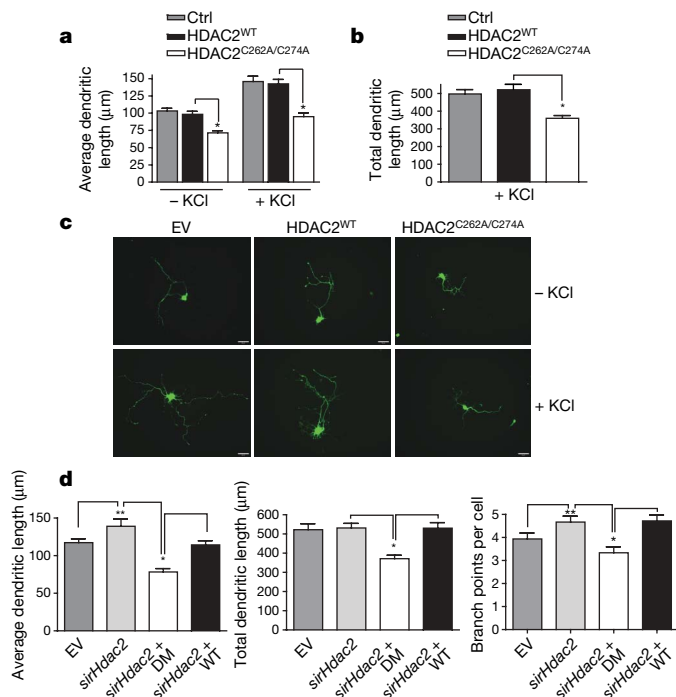


Figure 4 | Neurotrophin-dependent S-nitrosylation of HDAC2 regulates dendritic growth. **a**, **b**, Average (**a**) and total (**b**) dendritic length of cortical neurons transfected with HDAC2 vectors and grown under basal conditions (**a**) or stimulated with 50 mM KCl (**b**). At least 30 neurons were analysed for each condition for each experiment. Shown are the averages and the s.e.m.; $n = 3$; asterisk, $P = 0.001$. **c**, Examples of cortical neurons transfected with HDAC2 vectors and either stimulated with KCl or left untreated. Scale bar, 40 μ m. **d**, Dendritic growth analysis of neurons transfected with rat *sirHdac2* and empty vector, rat *sirHdac2* and mouse HDAC2^{WT} (WT), rat *sirHdac2* and mouse double mutant HDAC2^{C262A/C274A} (DM), or with empty vector alone (EV). The average dendritic length (left), the total dendritic length (middle) and the number of branch points per cell (right) were assessed. Shown are the averages and the s.e.m.; $n = 3$; asterisk, $P = 0.001$; double asterisk, $P < 0.05$.

dendritic growth, quantitative analysis was performed in cortical neurons maintained with 50 mM potassium chloride, a condition that induces neuronal depolarization and calcium influx²⁷. Notably, calcium-dependent activation of CREB target genes mediates the dendritic growth observed under these conditions²⁷. As previously shown, stimulation of cortical neurons with potassium chloride for 48 h significantly increased dendritic growth (Fig. 4a). Expression of HDAC2^{C262/274A}, but not HDAC2^{WT} or empty vector, reduced potassium-chloride-dependent axon growth and markedly affected neuronal morphology (Fig. 4a–c and Supplementary Fig. 7a). As an alternative approach, we inhibited HDAC2 by transfecting cortical cultures with a rat-specific HDAC2 short interfering RNA (siRNA; *sirHdac2*; Supplementary Fig. 7b). Cortical neurons were transfected with *sirHdac2* and mouse HDAC2^{WT}, mouse HDAC2^{C262A/C274A} or empty vector alone, and the average dendritic length, the total dendritic length and the number of branches per cell was assessed. Silencing of HDAC2 increased the average dendritic length and the number of branch points per neuron (Fig. 4d), an effect that was rescued by mouse HDAC2^{WT} expression. In contrast, HDAC2^{C262A/C274A} expression inhibited dendritic growth and the number of branch points per cell (Fig. 4d). Taken together, these findings indicate that HDAC2 S-nitrosylation regulates dendritic growth and branching, possibly by the activation of CREB-dependent genes.

In unstimulated neurons, HDAC2 is tightly associated with gene promoters and represses transcription by deacetylating histones and possibly other chromatin-associated proteins. After neurotrophin stimulation and membrane depolarization, NO is synthesized and can diffuse into the nucleus where it nitrosylates HDAC2, facilitating the dissociation of the enzyme from chromatin. The release of HDAC2 in repressor complexes, concomitant with the recruitment of nuclear activators (as histone acetyltransferases, for example), promotes histone acetylation and the onset of transcription (Supplementary Fig. 7c). nNOS is mainly responsible for NO production in the nervous system, and its expression is tightly regulated both temporally and spatially²⁸. In the developing brain, nNOS reaches the highest expression level between embryonic day 16 (E16) and postnatal day 0 (P0), and this expression corresponds with the migration of neuronal precursors from the ventricular zone to the external layers of the cortex. nNOS is also expressed in sensory neurons between E12 and E15, when these neurons are highly dependent on neurotrophins for survival and growth²⁸. Here we have shown that BDNF, NGF and possibly other neurotrophins, can influence chromatin remodelling in neurons by S-nitrosylation of HDAC2. It is therefore conceivable that NO mediates the epigenetic changes that are induced by neurotrophins and by synaptic activity, and determines the expression of genes responsible for cell-fate choice and neuronal migration. Our findings have further implications beyond the mechanisms that regulate the early development of the nervous system and may extend to various physiological functions in the adult brain. Exposing animals to a new and enriched environment, for example, can partially revert memory loss and neurodegeneration by mechanisms that involve changes in histone acetylation, and possibly NO signalling^{2,29}. Notably, in a mouse model of neurodegeneration, administration of the general HDAC inhibitor sodium butyrate markedly inhibits neuronal loss and improves memory performances²⁹. In such neurodegenerative disorders, manipulation of NO signalling may prove to be a more specific therapeutic approach than the use of general HDAC inhibitors.

METHODS SUMMARY

DAF-FM DA staining. Neurons were grown *in vitro* for 5 days and incubated with DMEM containing DAF-FM DA (Molecular Probes; 10 μ M for 30 min at 37 °C). Cells were treated with BDNF (75 ng ml⁻¹), NPA (300 μ M), BDNF in the presence of NPA, or BDNF after the removal of NPA. Neurons were stained with CellTracker and Hoechst to visualize cell bodies and nuclei, fixed with 4% paraformaldehyde and images were captured with a Leica DM 2500 confocal microscope.

Biotin-switch assay. The biotin-switch assay is a sensitive assay that has been successfully and extensively used to detect S-nitrosylated proteins by specifically labelling nitrosylated cysteines with a biotin moiety^{12,15–17}. Biotinylated proteins can be easily detected by biotin western blot or streptavidin precipitation followed by western blotting. It is essential to perform the assay in the dark, as exposure of the samples to ultraviolet light may result in non-specific signals. We performed the biotin-switch assay as described^{13,14} with minor modifications. The assay is explained in detail in the Methods.

Analysis of dendritic growth. E17 cortical neurons were cultured on glass coverslips coated with laminin and poly-D-lysine, and transfected after 4 h *in vitro*. Transfections were carried out using 0.12 μ g of GFP vector, 1.2 μ g of HDAC2 vectors, 15 nM of rat *sirHdac2* oligos (Dharmacon) and 2 μ l of Lipofectamine 2000 (Invitrogen). Neurons were incubated for 4 h at 37 °C and culture medium was replaced with the original plating medium with or without KCl (50 mM). Forty-eight hours after transfection, cells were fixed and quantitative analysis of dendritic growth and branching was assessed as described²⁷. Neurites were confirmed to be dendrites by MAP2 staining (A.N. and A.R., unpublished observations). Rat *Hdac2* siRNA sequences are available on request.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 23 May; accepted 7 July 2008.

Published online 27 August 2008.

- Huang, E. & Reichardt, L. Neurotrophins: roles in neuronal development and function. *Annu. Rev. Neurosci.* **24**, 677–736 (2001).
- Riccio, A. *et al.* A nitric oxide signaling pathway controls CREB-mediated gene expression in neurons. *Mol. Cell* **21**, 283–294 (2006).
- Reik, W. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**, 425–432 (2007).
- Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
- Huang, E. J. & Reichardt, L. F. Trk receptors: roles in neuronal signal transduction. *Annu. Rev. Biochem.* **72**, 609–642 (2003).
- Zweifel, L. S., Kuruvilla, R. & Ginty, D. D. Functions and mechanisms of retrograde neurotrophin signalling. *Nature Rev. Neurosci.* **6**, 615–625 (2005).
- Zhang, J. & Snyder, S. H. Nitric oxide in the nervous system. *Annu. Rev. Pharmacol. Toxicol.* **35**, 213–233 (1995).
- Boehning, D. & Snyder, S. H. Novel neural modulators. *Annu. Rev. Neurosci.* **26**, 105–131 (2003).
- Stamler, J. S., Lamas, S. & Fang, F. C. Nitrosylation: the prototypic redox-based signaling mechanism. *Cell* **106**, 675–683 (2001).
- Hess, D. T., Matsumoto, A., Kim, S. O., Marshall, H. E. & Stamler, J. S. Protein S-nitrosylation: purview and parameters. *Nature Rev. Mol. Cell Biol.* **6**, 150–166 (2005).
- Marshall, H. E., Hess, D. T. & Stamler, J. S. S-nitrosylation: Physiological regulation of NF- κ B. *Proc. Natl Acad. Sci. USA* **101**, 8841–8842 (2004).
- Stroissnigg, H. *et al.* S-nitrosylation of microtubule-associated protein 1B mediates nitric-oxide-induced axon retraction. *Nature Cell Biol.* **9**, 1035–1045 (2007).
- Jaffrey, S. R. & Snyder, S. H. The biotin switch method for the detection of S-nitrosylated proteins. *Sci. STKE* **2001**, PL1 (2001).
- Forrester, M. T., Foster, M. W. & Stamler, J. S. Assessment and application of the biotin switch technique for examining protein S-nitrosylation under conditions of pharmacologically induced oxidative stress. *J. Biol. Chem.* **282**, 13977–13983 (2007).
- Whalen, E. J. *et al.* Regulation of β -adrenergic receptor signaling by S-nitrosylation of G-protein-coupled receptor kinase 2. *Cell* **129**, 511–522 (2007).
- Uehara, T. *et al.* S-Nitrosylated protein-disulphide isomerase links protein misfolding to neurodegeneration. *Nature* **441**, 513–517 (2006).
- Hara, M. R. *et al.* S-nitrosylated GAPDH initiates apoptotic cell death by nuclear translocation following Siah1 binding. *Nature Cell Biol.* **7**, 665–674 (2005).
- Hausladen, A., Privalle, C. T., Keng, T., DeAngelo, J. & Stamler, J. S. Nitrosative stress: activation of the transcription factor OxyR. *Cell* **86**, 719–729 (1996).
- Cheung, P., Allis, C. D. & Sassone-Corsi, P. Signaling to chromatin through histone modifications. *Cell* **103**, 263–271 (2000).
- Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
- Wink, D. A. *et al.* Detection of S-nitrosothiols by fluorometric and colorimetric methods. *Methods Enzymol.* **301**, 201–211 (1999).
- Haendeler, J. *et al.* Redox regulatory and anti-apoptotic functions of thioredoxin depend on S-nitrosylation at cysteine 69. *Nature Cell Biol.* **4**, 743–749 (2002).
- Hassig, C. A. *et al.* A role for histone deacetylase activity in HDAC1-mediated transcriptional repression. *Proc. Natl Acad. Sci. USA* **95**, 3519–3524 (1998).
- Gu, Z. *et al.* S-nitrosylation of matrix metalloproteinases: signaling pathway to neuronal cell death. *Science* **297**, 1186–1190 (2002).
- Mustafa, A. K. *et al.* Nitric oxide S-nitrosylates serine racemase, mediating feedback inhibition of D-serine formation. *Proc. Natl Acad. Sci. USA* **104**, 2950–2955 (2007).

26. Huang, Y. *et al.* S-nitrosylation of N-ethylmaleimide sensitive factor mediates surface expression of AMPA receptors. *Neuron* **46**, 533–540 (2005).
27. Redmond, L., Kashani, A. H. & Ghosh, A. Calcium regulation of dendritic growth via CaM kinase IV and CREB-mediated transcription. *Neuron* **34**, 999–1010 (2002).
28. Bredt, D. S. & Snyder, S. H. Transient nitric oxide synthase neurons in embryonic cerebral cortical plate, sensory ganglia, and olfactory epithelium. *Neuron* **13**, 301–313 (1994).
29. Fischer, A., Sananbenesi, F., Wang, X., Dobbin, M. & Tsai, L. H. Recovery of learning and memory is associated with chromatin remodelling. *Nature* **447**, 178–182 (2007).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to D. D. Ginty, A. Lloyd, M. Raff and A. Saiardi for comments and to all members of the Riccio laboratory for discussion. We thank E. Seto for providing HDAC plasmids and A. Chittka for help with the purification of

recombinant HDAC2. We also thank C. Andreassi for help with cell lines and molecular biology techniques. This work was supported by the Medical Research Council (MRC; grant G0500792), and the European Research Council (Marie Curie International Reintegration grant MIRG-CT-2005-016501). A.R. is a recipient of an MRC Career Development Fellowship (G117/533) and P.M.W. of an MRC Career Development Award.

Author Contributions A.N. performed most of the experiments, analysed the data and helped to write the manuscript. P.M.W. performed the HDAC activity assay and some of the biotin-switch and ChIP assays. J.D.R. provided the DAF-FM DA data and contributed to the analysis of dendritic growth. L.C. designed and tested the siRNAs and helped with the HDAC2 constructs. A.R., the senior author, designed the project, performed most of the ChIP assays, analysed the data, wrote the manuscript and provided financial support.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.R. (a.riccio@ucl.ac.uk).

METHODS

Reagents. All biochemical reagents were purchased from Sigma and tissue culture reagents were purchased from Gibco (Invitrogen), unless otherwise stated. S-nitrosocysteine (SNOC) was prepared as described previously³⁰ and used within 1 h.

Cell cultures. Embryonic cortical rat and mouse neurons were cultured as previously described². Before experiments, cells were placed in low serum medium (3% fetal bovine serum, FBS) containing DL-2-amino-5-phosphonovaleric acid (DL-AP5, 50 μ M) for 12–16 h. HEK293T cells and PC12 cells were maintained in DMEM supplemented with 10% FBS, and 10% FBS plus 5% horse serum, respectively. Cells were transfected in OptiMem containing 20 μ g of DNA and 20 μ l of Lipofectamine 2000 (Invitrogen) and incubated for 4–6 h before replacing the transfection medium with serum-containing media. Before experiments, PC12 cells were starved in medium containing 0.5% horse serum overnight.

Biotin-switch assay. This assay was performed in the dark. Cells were lysed in HEN buffer (250 mM HEPES, 1 mM EDTA and 100 mM neocuproine) and adjusted to contain 0.4% CHAPS. Samples were homogenized and free cysteines were blocked for 1 h at 50 °C in three volumes of blocking buffer (HEN buffer plus 2.5% SDS, HENS) containing methyl methanethiosulfonate (200 mM). Proteins were acetone precipitated at –20 °C and resuspended in 100 μ l HENS solution. After adding fresh ascorbic acid (20 mM) and 1 mM biotin-HPDP (Pierce), proteins were incubated at room temperature for 1 h. After separation using an SDS–PAGE gel, biotinylated proteins were detected with streptavidin-HRP (Amersham). Alternatively, biotinylated proteins were resuspended in 100 μ l HENS buffer containing 200 μ l of neutralization buffer (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, 0.5% Triton X-100) and precipitated with 50 μ l of streptavidin–agarose beads at room temperature for 1 h. The beads were washed five times at 4 °C using neutralization buffer containing 600 mM NaCl. Biotinylated proteins were eluted using 50 μ l elution buffer (20 mM HEPES, 100 mM NaCl, 1 mM EDTA, 100 mM β -mercaptoethanol) and heated at 95 °C for 5 min in reducing SDS–PAGE loading buffer.

Extraction of nuclear and cytoplasmic proteins. Nuclear and cytoplasmic extracts were prepared using NE-PER nuclear and cytoplasmic extraction kit (Pierce) according to the manufacturer's instructions.

Chromatin immunoprecipitation. ChIP assays were performed mostly as described². For all cell types, approximately $3\text{--}5 \times 10^6$ cells were used per ChIP. In brief, medium was removed from treated cells and replaced with PBS containing 1% formaldehyde. Cells were rinsed twice in PBS, and collected in collection buffer containing 100 mM Tris-HCl, pH 9.4, and 1 mM dithiothreitol. Cells were collected by centrifugation, rinsed with PBS and pellets were resuspended in lysis buffer containing 0.1% SDS, 0.5% Triton X-100, 20 mM Tris-HCl, pH 8.1, and 150 mM NaCl. Samples were sonicated with six 20-s pulses and a 10-s interpulse interval. Cell debris was removed by centrifugation, and supernatants were pre-cleared by incubation with Protein A–Sepharose beads (Amersham Biosciences) for 1 h at 4 °C. Beads were collected by centrifugation and supernatants were subjected to immunoprecipitation. A fraction of the supernatant was used for immunoprecipitation input control. The volume of each tube was adjusted to 500 μ l with lysis buffer, and 5–10 μ g of rabbit polyclonal antibody was added overnight at 4 °C. The following polyclonal antibodies were used: anti-HDAC2 (Santa Cruz), anti-acetyl histone H3 K9/K14 (Upstate), anti-acetyl histone H4 (Upstate), control IgG (Santa Cruz) and anti-Flag M2 monoclonal antibody (Sigma). Immune complexes were collected by incubation with Protein A–Sepharose beads for 1 h at 4 °C. Beads were collected and subjected to a series of seven sequential washes, as described². Immune complexes were eluted from beads by vortexing in elution buffer containing 1% SDS and 0.1 M NaHCO₃. NaCl was added (final concentration 0.33 M), and crosslinking was reversed by incubation overnight at 65 °C. DNA fragments were purified using the QIAquick PCR purification kit (Qiagen). For PCR, specific sets of primers were designed that flank CRE regions within the upstream regulatory regions of the indicated genes. PCR conditions and cycle numbers were determined empirically for the different templates and primer pairs. Primers amplified fragments ranging in size from 200–400 bp. Primer sequences and PCR conditions are available on request.

Quantitative real-time PCR. PCR reactions (25 μ l) contained 12.5 μ l of PCR Sybr Green mix (NEB) and 0.3 μ M of primers. All reactions were performed in duplicate with an Opticon 2 System (MJ Research) and each experiment included a standard curve, a preimmune control and a NO template control. Standard templates consisted of gel-purified PCR products of the *Fos*, *Egr1*, *Vgf* and *Nos1* amplicons of known concentration, and each standard curve consisted of seven serial dilutions of the template. At the end of the 46 cycles of amplification, a dissociation curve was performed in which Sybr Green was measured at

1 °C intervals between 50 °C and 100 °C. Melting temperatures for *Fos*, *Egr1*, *Vgf* and *Nos1* were 83 °C, 88 °C, 92 °C and 85 °C, respectively. Results were normalized using total input DNA and expressed as fold changes over unstimulated control.

Purification of recombinant HDAC2. XL-1 Blue bacteria were transformed using a pGEX-HDAC2 construct. Cells were centrifuged and washed using a high salt buffer (20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 1 mM EDTA and protease inhibitors). Cells were resuspended in high salt buffer containing lysozyme (100 μ g ml^{–1}) and incubated at room temperature for 15 min. Dithiothreitol (5 mM) was added to cell lysates and after sonication, the lysates were clarified by adding DNaseI (2.5 μ g ml^{–1}, 15 min at room temperature). Glutathione beads were added and lysates were rotated for 30 min at 4 °C. Beads were washed twice with buffer A (20 mM Tris-HCl, pH 8.0, 200 mM NaCl and 10% glycerol), once with buffer B (20 mM Tris-HCl, pH 8.0, 0.5 M NaCl and 10% glycerol), once again with buffer A, and equilibrated with 50 mM Tris-HCl, pH 9.5. Proteins were eluted with 20 mM reduced glutathione in 50 mM Tris-HCl, pH 9.5. Fractions containing proteins were dialysed in PBS.

***Nos1* mouse line.** *Nos1* homozygous null animals³¹ or wild-type control animals were intercrossed to generate null or control litters, respectively. Embryos were collected at E16.5 for preparation of primary cortical cultures.

HDAC2 site-directed mutagenesis and truncations. The pME18S–HDAC2 (Flag-tagged) construct was a gift from E. Seto³². Point mutations were carried out using the Quickchange II site-directed mutagenesis kit (Stratagene) according to the manufacturer's instructions. The HDAC2 deletion constructs were generated using primers designed with 5' SalI and 3' NotI restriction sites to allow cloning in the pCMV–Myc plasmid (Clontech). The primers used for HDAC2 mutagenesis and deletions are available on request.

Immunoprecipitation. Lysates were prepared in immunoprecipitation buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP40) and sonicated using a BioRuptor (Diagenode). Lysates were pre-cleared with 50 μ l of Protein A–Sepharose beads (Amersham) and incubated with 5 μ g of primary antibody overnight on a tube rotator at 4 °C. The antibodies used for immunoprecipitation were: anti-HDAC2 rabbit polyclonal antibody (Santa Cruz) and anti-Flag M2 mouse monoclonal antibody (Sigma). Beads were washed five times in immunoprecipitation buffer and either boiled in 2 \times reducing SDS–PAGE loading buffer for western blotting or used in the HDAC assay.

Histone deacetylase assay. HDAC activity in whole cell lysates and immunoprecipitates was assessed using a fluorogenic HDAC assay³³. HDAC acts on the non-fluorescent substrate peptide BOC-(acetyl)Lys-AMC (Bachem), allowing subsequent trypsin digestion of the deacetylated peptide to yield a fluorescent moiety. For whole cell lysates, 15 μ l of extract was mixed with a 100 μ M fluorogenic HDAC peptide in HDAC buffer (25 mM Tris-HCl, pH 8.0, 137 mM NaCl, 2.7 mM KCl, 1 mM MgCl₂) to a final volume of 40 μ l. After incubation at 37 °C for 1 h, the reaction was stopped and the peptide was digested by incubation with 1 μ M trichostatin A (Tocris Bioscience) and 5 mg ml^{–1} trypsin (Sigma). The extracts were transferred to a black 96-well plate and read using a Spectramax Gemini plate-reader (Molecular Devices) at an excitation wavelength of 355 nm and an emission wavelength of 460 nm. For HDAC assay using immunoprecipitates, immunoprecipitated protein bound to Sepharose A beads was washed three times in immunoprecipitation lysis buffer and twice in HDAC buffer. Fluorogenic HDAC substrate was added to the beads at a final concentration of 100 μ M in 60 μ l HDAC buffer and incubated at 37 °C for 1 h, after which the reaction was stopped and fluorescence was measured as described above.

Dual luciferase assay. E17 cortical neurons were cultured in 6-well plates and transfected after 4 days *in vitro*. Transfections were carried out using 2 μ g of Fos–luciferase vector (firefly), 1 μ g of thymidine-kinase–luciferase (*Renilla*), 2 μ g of HDAC2 vectors and 10 μ l of Lipofectamine 2000 (Invitrogen). Neurons were incubated for 4 h at 37 °C and culture medium was replaced with the original plating medium. Forty-eight hours after transfection, cells were stimulated with 75 ng ml^{–1} BDNF, 200 μ M SNOC, 300 μ M NPA or NPA plus BDNF and samples were processed using the dual-luciferase reporter assay system (Promega), according to manufacturer's instructions.

30. Mallis, R. J., Buss, J. E. & Thomas, J. A. Oxidative modification of H-ras: S-thiolation and S-nitrosylation of reactive cysteines. *Biochem. J.* **355**, 145–153 (2001).
31. Huang, P. L., Dawson, T. M., Bredt, D. S., Snyder, S. H. & Fishman, M. C. Targeted disruption of the neuronal nitric oxide synthase gene. *Cell* **75**, 1273–1286 (1993).
32. Laherty, C. D. *et al.* Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell* **89**, 349–356 (1997).
33. Wegener, D., Wirsching, F., Riester, D. & Schwienhorst, A. A fluorogenic histone deacetylase assay well suited for high-throughput activity screening. *Chem. Biol.* **10**, 61–68 (2003).

LETTERS

Structure of the 30S translation initiation complex

Angelita Simonetti^{1,2,3,4}, Stefano Marzi^{1,2,3,4,5}, Alexander G. Myasnikov^{1,2,3,4}, Attilio Fabbretti⁶, Marat Yusupov^{1,2,3,4}, Claudio O. Gualerzi⁶ & Bruno P. Klaholz^{1,2,3,4}

Translation initiation, the rate-limiting step of the universal process of protein synthesis, proceeds through sequential, tightly regulated steps. In bacteria, the correct messenger RNA start site and the reading frame are selected when, with the help of initiation factors IF1, IF2 and IF3, the initiation codon is decoded in the peptidyl site of the 30S ribosomal subunit by the fMet-tRNA^{fMet} anticodon. This yields a 30S initiation complex (30SIC) that is an intermediate in the formation of the 70S initiation complex (70SIC) that occurs on joining of the 50S ribosomal subunit to the 30SIC and release of the initiation factors^{1–3}. The localization of IF2 in the 30SIC has proved to be difficult so far using biochemical approaches, but could now be addressed using cryo-electron microscopy and advanced particle separation techniques on the basis of three-dimensional statistical analysis. Here we report the direct visualization of a 30SIC containing mRNA, fMet-tRNA^{fMet} and initiation factors IF1 and GTP-bound IF2. We demonstrate that the fMet-tRNA^{fMet} is held in a characteristic and precise position and conformation by two interactions that contribute to the formation of a stable complex: one involves the transfer RNA decoding stem which is buried in the 30S peptidyl site, and the other occurs between the carboxy-terminal domain of IF2 and the tRNA acceptor end. The structure provides insights into the mechanism of 70SIC assembly and rationalizes the rapid activation of GTP hydrolysis triggered on 30SIC–50S joining^{2,3} by showing that the GTP-binding domain of IF2 would directly face the GTPase-activated centre of the 50S subunit.

The 30SIC complex from *Thermus thermophilus* was formed by incubating purified 30S ribosomal subunits, mRNA, initiator tRNA (fMet-tRNA^{fMet}), IF1 and IF2 in the presence of GTP which stabilizes the IF2–30S interactions⁴ (see Methods and ref. 5). The use of the natural IF2–ligand GTP rather than a non-hydrolysable analogue was possible because GTP remains stable in the 30SIC until the GTPase activity of IF2 is triggered on joining of the 50S subunit¹. The complexes prepared with the 30S subunits showed a higher tendency to aggregate than 70S ribosomes; this difficulty could be handled by careful particle picking and processing. Despite the optimization of the conditions to obtain the 30SIC (see ref. 5), during both structure determination by cryo-electron microscopy (cryo-EM) and three-dimensional reconstruction these complexes were still of limited homogeneity, a typical feature of large, multi-component complexes endowed in the present case with the intrinsic flexibility of the 30S ribosomal subunit. This problem was overcome by image processing and particle separation techniques on the basis of three-dimensional statistical analysis (see Methods), a method we developed that allows the study of several states (B.P.K., submitted) in contrast to that used previously⁶ which is limited to two states. From a single sample this procedure yielded several sets of three-dimensional structures (clustered to five statistically relevant structures), refined to a resolution of around 9 Å (Fig. 1 and Supplementary Fig. 1), that differ at the level of

composition or conformations. Whereas both IF1 and fMet-tRNA^{fMet} are visible in all sets, IF2 is absent in sets 4 and 5. A clear density for the helix formed by the Shine–Dalgarno/anti-Shine–Dalgarno base pairs is seen in all complexes except for sets 3 and 4. The set containing the full and best resolved complex (set 1) is described in detail and compared with the other sets.

Comparison of the 30SIC cryo-EM structure with that of a control 30S mRNA complex (see Methods), with the *T. thermophilus* 30S crystal structures^{7,8} and with the 30S subunit of the *T. thermophilus* 70S crystal structure^{9,10}, demonstrates a stretched-out density on the 30S surface that would face the 50S subunit in a 70SIC (see Fig. 2a). This density extends from the helix h14 (16S RNA; ‘h’ and ‘H’ are used to number rRNA helices of the small and large subunit, respectively) area to the tRNA peptidyl (P)-site area and is attributable to IF2 and fMet-tRNA^{fMet}. Aside from some important functional differences discussed later, the positions of IF2 and of fMet-tRNA^{fMet} correspond to those seen for other ribosome-dependent GTPases bound to the factor binding-site and to the 70S P-site-bound tRNA, respectively. An additional small density is seen close to the decoding site; from its size and position it is attributable to IF1 exactly as observed by soaking IF1 into 30S crystals¹¹. To analyse the molecular interactions within the 30SIC, structures of the individual components were fitted to the experimental map. For this, we used the crystal structure of the *T. thermophilus* 30S–IF1 complex¹¹, the P-site tRNA of the *T. thermophilus* 70S crystal structures¹⁰ and the *T. thermophilus* IF2 homology model derived from the archaeal aIF5B crystal structure^{12,13}. The map shows two large structural modules for IF2 connected by a narrow density; the size of these modules suggests that they correspond to sub-domains I/II (G2/G3) and III/IV (C1/C2). Accordingly, the 30S-bound bacterial IF2 seems to be in a new conformation in which domain III (C1) is shifted towards domain IV (C2). The amino-terminal and G1 domains are not seen in the maps, suggesting that this part of IF2 is disordered in the 30SIC.

The molecular model obtained shows two anchor points for the IF2–fMet-tRNA^{fMet} sub-complex on the 30S subunit: one through IF2 and one through the tRNA decoding stem, thus explaining the stabilization of the fMet-tRNA^{fMet} by IF2 on the 30S subunit even in the absence of the 50S subunit. Domains I/II (G2/G3) are in the vicinity of helices h5 and h14 of the 16S RNA (Supplementary Fig. 2). The second anchor point is about 95 Å away, where the decoding stem of the initiator fMet-tRNA^{fMet} is buried in a pocket provided by the neck of the 30S ribosomal subunit as seen in the *T. thermophilus* 70S ribosome crystal structures with bound mRNA and tRNA^{10,14}. On the side oriented towards the aminoacyl (A)-site, the tRNA stem interacts with ribosomal protein S13; on the other side, the tRNA stem interacts with helix h24 and the extended C-terminal tail of protein S9.

The bridge formed by IF2 and tRNA that connects the two anchor points on the 30S subunit is maintained by a tight interaction (strong continuous density in the cryo-EM map) between the C-terminal

¹Institute of Genetics and of Molecular and Cellular Biology, Department of Structural Biology and Genomics, Illkirch F-67404, France. ²Inserm, U596, Illkirch F-67404, France. ³CNRS, UMR7104, Illkirch F-67404, France. ⁴Université Louis Pasteur, Strasbourg F-67070, France. ⁵Architecture et Réactivité de l'ARN, UPR 9002 CNRS, Institute of Molecular and Cellular Biology, 15 rue R. Descartes, 67084 Strasbourg, France. ⁶Laboratory of Genetics, Department of Biology MCA, University of Camerino, 62032 Camerino (MC), Italy.

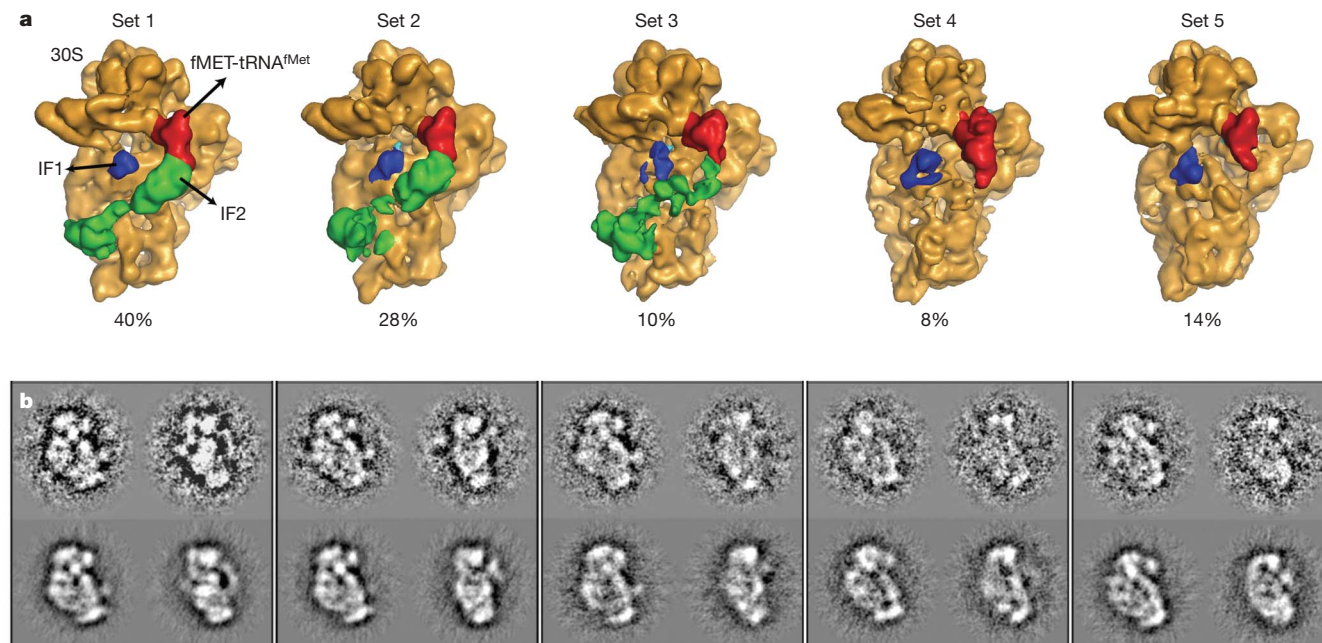


Figure 1 | One sample—several structures. Several states of the bacterial 30SIC from a single sample visualized by cryo-EM and advanced particle separation on the basis of three-dimensional statistical analysis. The percentage of particles forming each set is given (total data set size: 80,285 single particle images). **a**, The density parts corresponding to 30S, fMet-tRNA^{fMet}, IF1 and IF2 are coloured in orange, red, blue and green, respectively. The structure subsets of the 30SIC differ in conformation and

composition: IF1 and fMet-tRNA^{fMet} are present in all sets, albeit the conformation of fMet-tRNA^{fMet} is different in the last two sets in which IF2 is not present; the mRNA is not seen in sets 3 and 4 (see text for further details). **b**, Two pairs of representative class averages and the corresponding re-projections of the three-dimensional reconstruction for each structure subset.

domain IV (C2) of IF2, the conserved 3' hexa-nucleotide CAACCA end^{15,16} and part of the double-stranded region of the tRNA (see Fig. 2b). Contouring the cryo-EM map at different levels (see insert in Fig. 2b) shows that the 3' tail of the acceptor stem is kinked around the position of residues C72–A73, resulting in the 3'-adenine being positioned about 15 Å away from that of a linear CCA tail. This conformational change is induced by IF2 and increases the interface between the fMet-tRNA^{fMet} and the factor. Compared to the standard tRNA structures in both free and 70S-bound states, the fitted tRNA also shows a distortion at the level of the decoding stem which is bent towards the initiation codon. A similar distortion has been seen for A-/T-site tRNA bound to the 70S in the presence of elongation factor EF-Tu¹⁷ and, to a lesser extent, for the P-site tRNA in the 70S crystal structure⁹. The observed change in the fMet-tRNA^{fMet} structure apparently depends on the presence of both codon–anticodon interactions and of IF2, because the bend is not seen in the structures of sets 4 (when IF2 and mRNA are absent) and 5 (when IF2 is absent; Fig. 1).

The fMet-tRNA^{fMet} position is rather different from that of the classical P-site tRNA observed in 70S crystal structures¹⁰ or in the 70S–IF2 complex analysed by cryo-EM¹³. Whereas the anticodon loop is essentially in the 30S P-site, the anticodon stem is slightly rotated clockwise in the P-site pocket and the tRNA elbow is shifted towards the exit (E)-site. As a result, the fMet-tRNA^{fMet} is tilted, the elbow and acceptor stem lie between P- and E-sites (Fig. 3a), and the CCA tail is lifted ~15 Å away from the peptidyl-transferase centre position observed in 70S–tRNA complexes. Intermediate positions of tRNAs have been so far described during translocation when A/P and P/E states are transiently adopted in 70S ribosomes¹⁸. The present discovery of tRNA intermediate states in the 30SIC suggests that accommodating the tRNA in transient states is an important feature of the small subunit and of its flexible nature. During translocation in the 70S ribosome, the three tRNAs remain positioned within a plane (Fig. 3a), whereas the fMet-tRNA^{fMet} seen in a 70SIC containing IF1 and IF2 moves out of the 'translocation plane' into a position called

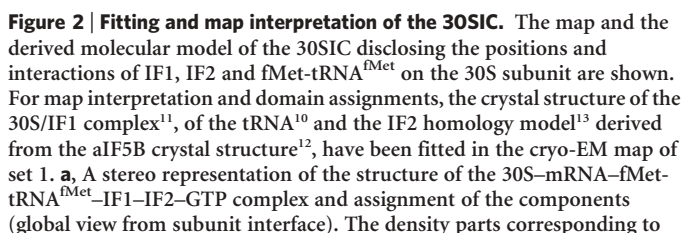
the peptidyl/initiation (P/I) state¹⁹. In the present 30SIC, the fMet-tRNA^{fMet} adopts a position (we refer to as the 30S P/I state) related to, but yet different from, that of the 70SIC–IF1–IF2 complex: in fact, the fMet-tRNA^{fMet} remains in-plane with the classical tRNAs but its elbow is shifted towards helix h24 of the 16S ribosomal RNA. The position of the fMet-tRNA^{fMet} in the 30SIC is such that in the 70SIC it would face the 50S subunit between the P- and the E-sites. In contrast, in the 70SIC it is accommodated in the P/P-site, indicating that a tRNA conformational change occurs on subunit association.

In addition to alterations to the fMet-tRNA^{fMet}, the conformation of IF2 also changes during the initiation process (Fig. 3b). The comparison of the present structure of the 30S–IF1–IF2–GTP complex with that of the 70S–IF2–GMPPCP and 70S–IF2–GDP complexes¹³, which correspond to the states occurring before and after P_i dissociation, respectively, shows the existence of a conformation adopted by IF2 after subunit joining but before P_i release, and of another conformation adopted by the factor after P_i release (Fig. 3b). The existence of several discrete conformational changes of IF2 is consistent with results obtained by fast kinetics experiments, in which such changes are observed on GTP hydrolysis and P_i release in real time by fluorophore-labelled IF2 and tRNA³.

The observed conformational changes of both fMet-tRNA^{fMet} and IF2 provide insights into the subunit joining mechanism (Fig. 3c). The structural comparison of the 30SIC with 70S ribosome complexes^{9,10,13,14} shows that IF2 domain I/II (G2/G3) would clash with the α-sarcosine-ricin loop region of the 50S ribosomal subunit, and the fMet-tRNA^{fMet} would be close to loop H81/H83 (located between the P- and E-site tRNA binding pockets on the 50S subunit). It is therefore probable that in the early phase of subunit joining, the exposed G domain is contacted by the GTPase-associated centre of the 50S subunit. This mechanism is consistent with the fact that GTP hydrolysis occurs immediately on subunit joining, as shown previously by fast kinetic studies^{2,3}. Similarly, the particularly exposed position of the fMet-tRNA^{fMet} acceptor stem seems designed for interacting with the 50S subunit in the early phase of 70S assembly.

association²². Notably, the largest part of IF2 is complementary in shape to the 50S surface. In the 30SIC, the tRNA position is different from its final position in the 70SIC, suggesting that the tRNA transmits signals to IF2, and vice versa. In particular, after P_i release from IF2 and dissociation of the IF2/CCA-end contacts, the fMet-tRNA^{fMet} acceptor stem will back-translocate into the P/P state ready for the formation of the first peptide bond, whereas IF2, having lost its main anchorage to the 70S ribosomes, would be ready to dissociate^{1–4,13}.

The architecture of the 30SIC complex containing fMet-tRNA^{fMet}, IF2 and IF1, offers several clues as to the function of these ribosomal ligands during the initiation of protein synthesis; IF2 and tRNA are shown to have cooperative roles for their mutual stabilization on the 30S and for their function during subunit joining. The study also provides the localization and structure of 30S-bound IF2, representing the structure of a ribosomal GTPase bound to the ribosome in its natural GTP state. Moreover, our study provides, to our knowledge, the first structural investigation of a functional tRNA complex of the



b, View of the IF2–fMet-tRNA^{fMet} sub-complex (view rotated to the left compared to **a**). Domain assignment of IF2 and analysis of its interaction with the fMet-tRNA^{fMet} is shown. The tRNA decoding stem is bent towards the mRNA (the anti-codon is shown in green; the residual density next to it corresponds to the AUG start codon shown in dark green). Notably, the 3' CCA end of the tRNA (insert, view from 50S side) is kinked towards IF2 domain IV (C2) compared to a classical tRNA (in yellow), as seen from differential protein/RNA density contouring levels (shown in pink and orange).

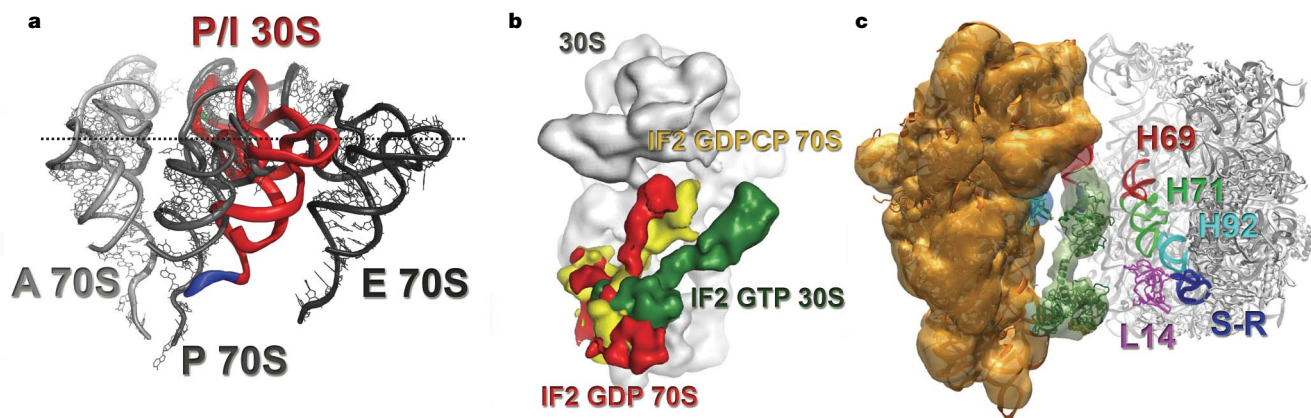


Figure 3 | Comparisons of the tRNA and IF2 positions in different translation initiation complexes, and model of ribosomal subunit joining. **a**, Comparison of the tRNA position in the 30SIC (30S–mRNA–fMet-tRNA^{fMet}–IF1–IF2–GTP complex, red) with those in the post-initiation complex 70S–mRNA–tRNA (Protein Data Bank accession 2HGP (ref. 10); A 70S for the A-site, P 70S for the P-site and E 70S for the E-site tRNAs, coloured in light grey, dark grey and black, respectively). Because the fMet-tRNA^{fMet} is held by IF2, it adopts an intermediate position (P/I 30S state) with the CCA tail (in blue) lifted out of the position corresponding to the peptidyl transferase centre on the 50S subunit. However, because it is also tilted towards the E-site, the fMet-tRNA^{fMet} elbow remains in the tRNA translocation plane (indicated by the dotted line). **b**, IF2 positions before and after joining of the ribosomal subunits¹³ representing the states before (IF2 GTP 30S, green) and after (IF2 GDP 70S, red) GTP hydrolysis, and

after P_i release (IF2 GDP 70S, red); the 30S model in the back is shown for general orientation. **c**, Model of the 50S subunit joining the 30S initiation complex, leading to the formation of the 70SIC. In the 30S–mRNA–fMet-tRNA^{fMet}–IF1–IF2 complex, the prominent position of the fMet-tRNA^{fMet} and of IF2 would provide first contact points with the 50S subunit. The exposed position of the IF2 G domain would contact the α -sarcin-ricin loop (S-R, blue; GTPase-activated centre) of the 23S rRNA, rationalizing that GTP hydrolysis occurs on subunit joining. IF2 and the 50S subunit have mostly complementary shapes, but some clashes can be expected with L14 (magenta), H92 (cyan) and H71 (green). Helix H69 (red) could slide underneath the elbow of fMet-tRNA^{fMet} without clashing with IF2 or the fMet-tRNA^{fMet}. IF2 and fMet-tRNA^{fMet} clearly require conformational changes such as those shown in **a** and **b** to adopt their positions in 70S initiation complexes.

30S ribosomal subunit. IF2 and IF1 are shown to not interact with each other on the 30S, in contrast to a widely accepted hypothesis; rather, the stimulation of the IF2 function by IF1 might be mediated by IF1-induced conformational changes of the 30S subunit^{11,23}. It should be noted that the structure of the 30SIC that is reported here does deliberately not contain IF3 to avoid a further increase of the conformational dynamics of the ribosomal subunit that this factor would have induced; productive 70SIC complexes can be formed also in the absence of this factor²⁴. Furthermore, the study shows that conformations, positions and domain arrangements of ligands are different in 30S and 70S initiation complexes, thus giving insights into the molecular mechanism of subunit joining, GTP hydrolysis and 70S formation, an important event in the regulation of cellular protein synthesis.

METHODS SUMMARY

The 30SIC complexes were formed by incubating 30S subunits from *T. thermophilus* with GTP, mRNA, fMet-tRNA^{fMet}, *T. thermophilus* IF2 and IF1 (see Methods and ref. 5) and applied to cryo-EM holey carbon grids as previously described⁶. 30S subunits containing only mRNA were used as controls. Low-dose images were collected on film with the in-house 200 kV FEI Tecnai20 and processed as described^{6,13,25,26}. There were 80,285 particle images used for the 30SIC, and 10,040 for the 30S–mRNA reference complex. As the 30SIC sample showed significant heterogeneity, a three-dimensional sampling and classification approach was applied to the 30SIC data set (B.P.K., submitted). The principle of the method is the generation of three-dimensional reconstructions from randomly selected particle images, and three-dimensional multivariate statistical analysis and three-dimensional classification into structural sets with low intra-class variance. The approach makes use of the principle in statistical analysis that resampling allows estimating the global (three-dimensional) variability from the variability of subensembles/subsamples²⁷; such resampling has been used in other cases (for example, ref. 28), in particular for localizing variable areas and sorting particles²⁹, and here in direct combination with three-dimensional multivariate statistical analysis and three-dimensional classification of structures. The number of sets is determined such that the sets show high inter-class variance together with low intra-class variance; the states thus represent structures with the highest variability and are representative of the structural states in the sample; independent validation is provided by the correlation between input

images and three-dimensional re-projections. The resolutions of the final three-dimensional structures were estimated with the Fourier shell correlation (see Supplementary Fig. 1) according to the 0.14 (ref. 30) and one-half-bit³¹ criteria to ~ 9 Å for the different sets of the 30SIC, and 20 Å for the 30S–mRNA reference complex.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 February; accepted 23 June 2008.

Published online 31 August 2008.

- Gualerzi, C. O. *et al.* Initiation factors in the early events of mRNA translation in bacteria. *Cold Spring Harb. Symp. Quant. Biol.* **66**, 363–376 (2001).
- Tomšić, J. *et al.* Late events of translation initiation in bacteria: a kinetic analysis. *EMBO J.* **19**, 2127–2136 (2000).
- Grigoriadou, C., Marzi, S., Kirillov, S., Gualerzi, C. O. & Cooperman, B. S. A quantitative kinetic scheme for 70S translation initiation complex formation. *J. Mol. Biol.* **373**, 562–572 (2007).
- Caserta, E. *et al.* Translation initiation factor IF2 interacts with the 30S ribosomal subunits via two separate binding sites. *J. Mol. Biol.* **362**, 787–799 (2006).
- Simonetti, A. *et al.*, *Nature Protocols*. doi:10.1038/nprot.2008.130 (2008).
- Klaholz, B. P., Myasnikov, A. G. & van Heel, M. Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature* **427**, 862–865 (2004).
- Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
- Schluenzen, F. *et al.* Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **102**, 615–623 (2000).
- Korostelev, A., Trakhanov, S., Laurberg, M. & Noller, H. F. Crystal structure of a 70S ribosome–tRNA complex reveals functional interactions and rearrangements. *Cell* **126**, 1065–1077 (2006).
- Yusupova, G., Jenner, L., Rees, B., Moras, D. & Yusupov, M. Structural basis for messenger RNA movement on the ribosome. *Nature* **444**, 391–394 (2006).
- Carter, A. P. *et al.* Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* **291**, 498–501 (2001).
- Roll-Mecak, A., Cao, C., Dever, T. E. & Burley, S. K. X-ray structures of the universal translation initiation factor IF2/elf5B: conformational changes on GDP and GTP binding complex. *Cell* **103**, 781–792 (2000).
- Myasnikov, A. G. *et al.* Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nature Struct. Mol. Biol.* **12**, 1145–1149 (2005).
- Selmer, M. *et al.* Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* **313**, 1935–1942 (2006).

15. Szkaradkiewicz, K., Zuleeg, S., Limmer, S. & Sprinzl, M. Interaction of fMet-tRNA^{fMet} and fMet-AMP with the C-terminal domain of *Thermus thermophilus* translation initiation factor 2. *Eur. J. Biochem.* **267**, 4290–4299 (2000).
16. Guenneugues, M. *et al.* Mapping the fMet-tRNA binding site of initiation factor IF2. *EMBO J.* **19**, 5233–5249 (2000).
17. Valle, M. *et al.* Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J.* **21**, 3557–3567 (2002).
18. Moazed, D. & Noller, H. F. Intermediate states in the movement of transfer RNA in the ribosome. *Nature* **342**, 142–148 (1989).
19. Allen, G. S., Zavialov, A., Gursky, R., Ehrenberg, M. & Frank, J. The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell* **121**, 703–712 (2005).
20. Gualerzi, C. O. & Pon, C. L. Initiation of mRNA translation in prokaryotes. *Biochemistry* **29**, 5881–5889 (1990).
21. Antoun, A., Pavlov, M. Y., Tenson, T. & Ehrenberg, M. Ribosome formation from subunits studied by stopped-flow and Rayleigh light scattering. *Biol. Proceed. Online* **6**, 35–54 (2004).
22. Ali, I. K., Lancaster, L., Feinberg, J., Joseph, S. & Noller, H. F. Deletion of a conserved, central ribosomal intersubunit RNA bridge. *Mol. Cell* **23**, 865–874 (2006).
23. Milon, P., Konevega, A. L., Gualerzi, C. O. & Rodnina, M. V. Kinetic checkpoint at a late step in translation initiation. *Mol. Cell* **30**, 712–720 (2008).
24. Giuliodori, A. M., Giangrossi, M., Brandi, A., Gualerzi, C. O. & Pon, C. L. Cold-stress-induced de novo expression of infC and role of IF3 in cold-shock translational bias. *RNA* **13**, 1355–1365 (2007).
25. Klaholz, B. P. *et al.* Structure of the *Escherichia coli* ribosomal termination complex with release factor 2. *Nature* **421**, 90–94 (2003).
26. Marzi, S. *et al.* Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. *Cell* **130**, 1019–1031 (2007).
27. Efron, B. Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. *Biometrika* **68**, 589–599 (1981).
28. Haynor, D. R. & Woods, S. D. Resampling estimates of precision in emission tomography. *IEEE Trans. Med. Imaging* **8**, 337–343 (1989).
29. Penczek, P. A., Yang, C., Frank, J. & Spahn, C. M. Estimation of variance in single-particle reconstruction using the bootstrap technique. *J. Struct. Biol.* **154**, 168–183 (2006).
30. Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).
31. van Heel, M. & Schatz, M. Fourier shell correlation threshold criteria. *J. Struct. Biol.* **151**, 250–262 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Thompson for providing IF2 strains, M. Argenti for mass spectroscopy analysis of IF1 and IF2, and P. Schultz, D. Moras, J.-C. Thierry, V. Mallouh, G. Yusupova and I. Orlov for their constant support and interest. This work was supported by grants from the Centre National pour la Recherche Scientifique (CNRS), the Ministère de la Recherche et de la Technologie, the European Molecular Biology Organization Young Investigator Programme, the Institut du Développement et des Ressources en Informatique Scientifique, and the European Commission as SPINE2-complexes (contract n° LSHG-CT-2006-031220). A.S. is a PhD student in a co-tutorial between the Université Louis Pasteur (ULP) and the Università di Camerino, and was supported by SPINE2-complexes, by the Institut National de la Santé et de la Recherche Médicale (INSERM) and by the Fondation de la Recherche Médicale (FRM). S.M. was supported by postdoctoral fellowships from the ULP, the CNRS and from the FRM, and A.G.M. was a recipient of postdoctoral fellowships from the CNRS and the FRM. The electron microscope facility is supported by the Alsace Region, the INSERM, the CNRS and the Association pour la Recherche sur le Cancer.

Author Information The electron density map of the 30SIC complex has been deposited to the EM Data Bank under the accession number EMD-1523. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to B.P.K. (klaholz@igbmc.u-strasbg.fr).

METHODS

Formation of 30S initiation complexes. Cloning, expression and purification of IF1 and IF2 from *T. thermophilus* were performed as described³² except that additional hydrophobic chromatography on a phenyl sepharose FF (Pharmacia) column was performed (more details can be found in ref. 5). The integrity of IF1 and IF2 has been confirmed by gel electrophoresis followed by MALDI peptides mass finger printing analysis. The fMet-tRNA^{fMet} was prepared according to ref. 33 using purified methionyl-tRNA synthetase and methionyl-tRNA^{fMet} formyl-transferase; 30S subunits from *T. thermophilus* were derived from 70S 'tight couples' ribosome, both purified according to ref. 34 by dialysis of the 70S in dissociation buffer (20 mM Tris-HCl, pH 7.5, 150 mM KCl, 1 mM MgAc₂ and 1 mM dithiothreitol) followed by ultracentrifugation on a 5%–20% sucrose gradient in the same buffer. The model 27-nucleotide mRNA (mk27) was from Dharmacon and purified as previously described³⁵; the GTP was from Roche Diagnostics. The 30SIC complexes, as well as the 30S/mRNA reference complex, were formed in 10 mM HEPES, pH 7.5, 70 mM NH₄Cl, 30 mM KCl, 8 mM MgAc₂ and 1 mM dithiothreitol by incubating 30S subunits from *T. thermophilus* (0.6 μM) with GTP (1 mM), mRNA (0.9 μM), fMet-tRNA^{fMet} (1.2 μM), *T. thermophilus* IF2 and IF1 (2.4 μM). After 20 min incubation at 50 °C the MgAc₂ concentration was increased to 20 mM at 20 °C. Occupancy of fMet-tRNA^{fMet} was 83%, as measured in a parallel reaction by a nitrocellulose filter binding assay using f[35S]Met-tRNA^{fMet} (see ref. 5); the fraction missing tRNA could have formed aggregates (as free 30S particles do) and were thus not used during particle selection, explaining why none of the structure subsets miss tRNA; alternatively, some tRNA may have become de-acylated during the 50 °C incubation and may therefore not interact with IF2 (such as sets 4 and 5).

Filter binding assays. To get more homogeneous particles for cryo-EM analysis, we optimized the formation of a stable and physiologically relevant 30SIC. For this purpose, the effect of temperature, ion concentration and the stoichiometry of initiation factors and mRNA⁵ on the 30SIC formation have been measured by nitrocellulose filter binding assays. Ribosomal occupancy of fMet-tRNA^{fMet} has been obtained using f[35S]Met-tRNA^{fMet}. Unless specified in ref. 5, the standard reaction mixture to form 30SIC was performed in a buffer containing 10 mM HEPES, pH 7.5, 30 mM KCl, 70 mM NH₄Cl, 8 mM MgAc₂ and 1 mM dithiothreitol. The concentration of f[35S]Met-tRNA^{fMet} was 0.9 μM and of 30S ribosomal subunits was 0.6 μM. After 20 min incubation at 50 °C, the reaction was stopped by the addition of 1 ml of ice-cooled MWB buffer containing 10 mM Tris-HCl, pH 7.1, 7 mM MgAc₂, 100 mM NH₄Cl and 2 mM β-mercaptoethanol, and filtered through nitrocellulose membrane (BA-85 filters; Schleicher & Schuell). The filters were washed twice with the same MWB buffer, dried and then the amount of f[35S]Met-tRNA/30SIC retained on the filters was determined by scintillation counting. Taken together, the results allowed us to set up

physical conditions, salts composition of the buffer and the stoichiometry of the individual ligands to be used to obtain the highest ribosome occupancy.

Cryo-EM structure determination and interpretation. 30SIC at a final concentration of 0.5 mg ml⁻¹ was diluted fourfold with the buffer used for complex assembly and 3.5 μl of the sample was applied to cryo-EM holey carbon grids as described previously⁶. Low-dose images were taken on Kodak SO-163 films with the in-house FEI Tecnai20 field emission gun transmission electron microscope using a magnification of ×50,000 and a defocus range of −0.8 to −3.0 μm. Data collection and processing were performed as described^{6,13,25,26}. The complexes prepared with the 30S subunits showed a higher tendency to aggregate than those made with 70S ribosomes; this difficulty could be handled by careful particle picking and processing. Initial structures showed strong heterogeneity. By using a three-dimensional sampling and classification approach (B.P.K., submitted), the 30SIC data set containing 80,285 images was split into 5 final subsets containing 32,000, 22,000, 8,000, 6,000 and 11,000 particles, respectively. The three-dimensional structures were calculated with IMAGIC-V³⁶ before particle sorting, and time-efficient three-dimensional reconstructions for three-dimensional multivariate statistical analysis and for the final structure refinement of the individual sets were done with the back projection algorithm BKPR³⁷ modified to comprise an exact weighting scheme applicable to all point group symmetries (Orlov *et al.*, in preparation). The high-frequency components of the final three-dimensional maps were suppressed by low-pass filtering at 9 Å. Manual fitting was performed in O using the 30S/IF1 crystal structure¹¹, the P-site tRNA of the *T. thermophilus* 70S crystal structure¹⁰ and the *T. thermophilus* IF2 homology model derived from the eIF5B crystal structure^{12,13}. The 30S head and body parts were fitted as separate rigid bodies and the figures were prepared using Pymol (<http://www.pymol.org>).

32. Thompson, J. & Dahlberg, A. E. Testing the conservation of the translational machinery over evolution in diverse environments: assaying *Thermus thermophilus* ribosomes and initiation factors in a coupled transcription-translation system from *Escherichia coli*. *Nucleic Acids Res.* **32**, 5954–5961 (2004).
33. Rodnina, M. V., Semenov, Y. P. & Wintermeyer, W. Purification of fMet-tRNA(fMet) by fast protein liquid chromatography. *Anal. Biochem.* **219**, 380–381 (1994).
34. Fechter, P. *et al.* Ribosomal initiation complexes probed by toeprinting and effect of trans-acting translational regulators in bacteria. *Methods Mol. Biol.* (in the press).
35. Yusupova, G. Z., Yusupov, M. M., Cate, J. H. & Noller, H. F. The path of the messenger RNA through the ribosome. *Cell* **106**, 233–241 (2001).
36. van Heel, M., Harauz, G., Orlova, E. V., Schmidt, R. & Schatz, M. A new generation of the IMAGIC image processing system. *J. Struct. Biol.* **116**, 17–24 (1996).
37. Orlov, I. M., Morgan, D. G. & Cheng, R. H. Efficient implementation of a filtered back-projection algorithm using a voxel-by-voxel approach. *J. Struct. Biol.* **154**, 287–296 (2006).

Prolyl 4-hydroxylation regulates Argonaute 2 stability

Hank H. Qi¹, Pat P. Ongusaha², Johanna Myllyharju³, Dongmei Cheng⁴, Outi Pakkanen³, Yujiang Shi⁵, Sam W. Lee², Junmin Peng⁴ & Yang Shi¹

Human Argonaute (Ago) proteins are essential components of the RNA-induced silencing complexes (RISCs). Argonaute 2 (Ago2) has a P-element-induced wimpy testis (PIWI) domain, which folds like RNase H and is responsible for target RNA cleavage in RNA interference¹. Proteins such as Dicer, TRBP, MOV10, RHA, RCK/p54 and KIAA1093 associate with Ago proteins and participate in small RNA processing, RISC loading and localization of Ago proteins in the cytoplasmic messenger RNA processing bodies^{1,2}. However, mechanisms that regulate RNA interference remain obscure. Here we report physical interactions between Ago2 and the α -(P4H- α (I)) and β -(P4H- β) subunits of the type I collagen prolyl-4-hydroxylase (C-P4H(I)). Mass spectrometric analysis identified hydroxylation of the endogenous Ago2 at proline 700. *In vitro*, both Ago2 and Ago4 seem to be more efficiently hydroxylated than Ago1 and Ago3 by recombinant human C-P4H(I). Importantly, human cells depleted of P4H- α (I) or P4H- β by short hairpin RNA and P4H- α (I) null mouse embryonic fibroblast cells showed reduced stability of Ago2 and impaired short interfering RNA programmed RISC activity. Furthermore, mutation of proline 700 to alanine also resulted in destabilization of Ago2, thus linking Ago2 P700 and hydroxylation at this residue to its stability regulation. These findings identify hydroxylation as a post-translational modification important for Ago2 stability and effective RNA interference.

To identify the protein network involved in regulating the RNA interference (RNAi) machinery, we established stable HeLa S3 cell lines expressing Flag-HA(haemagglutinin)-tagged human Ago1–4, respectively. Cytoplasmic extracts were immunoprecipitated with sequential anti-Flag and anti-HA antibody resins. Silver staining of a representative purification (Ago2) is shown in Fig. 1a. Mass spectrometric analysis of the purifications identified the known Ago-interacting proteins¹, among which Dicer, KIAA1093 and eEF1A1 are present in all four Ago purifications, whereas TRBP and MOV10 are abundantly present in Ago3 and 4 purifications. In addition, unique peptides of P4H- α (I) and P4H- β , the α - and β -subunits of C-P4H(I) (EC 1.14.11.2), were found in sub-stoichiometric quantities in Ago2 purification (Fig. 1a). To confirm the P4H- α (I) interaction with Ago proteins, we performed co-immunoprecipitation experiments using cell lysates from 293ET cells co-transfected with Flag-tagged Ago 1–4 and HA-tagged P4H- α (I), and found interactions of P4H- α (I) with Ago2 and Ago4 (Fig. 1b). Co-immunoprecipitation using a polyclonal anti-P4H- β antibody co-precipitated endogenous Ago2 from 293ET cells (Fig. 1c), supporting the *in vivo* interaction between C-P4H(I) subunits with Ago2. That a small percentage of Ago2 protein was found to associate with P4H- β may reflect the dynamic nature of this enzyme–substrate interaction.

C-P4H catalyses proline hydroxylation of collagens in the X-Pro-Gly (X-P-G) triplets. Three isoforms (I, II, III) of the α -subunit have been identified³ and they all interact with the β -subunit to form active C-P4H $\alpha_2\beta_2$ tetramers. However, these isoforms show differential tissue expression pattern and abundance. In addition to collagens, other proteins containing collagen-like sequences, such as pulmonary surfactant, apoproteins, complement protein C1q and the 18S form of acetylcholinesterase, are also hydroxylated by C-P4H (ref. 3). C-P4H also mediates hydroxylation of the X-Pro-Ala (X-P-A) motif found in the complement protein C1q and type IV collagen⁴. We identified both X-P-G and X-P-A triplets within the Ago proteins (Supplementary Fig. 1). Three X-P-G triplets (119–121, 522–524, 699–701) were found

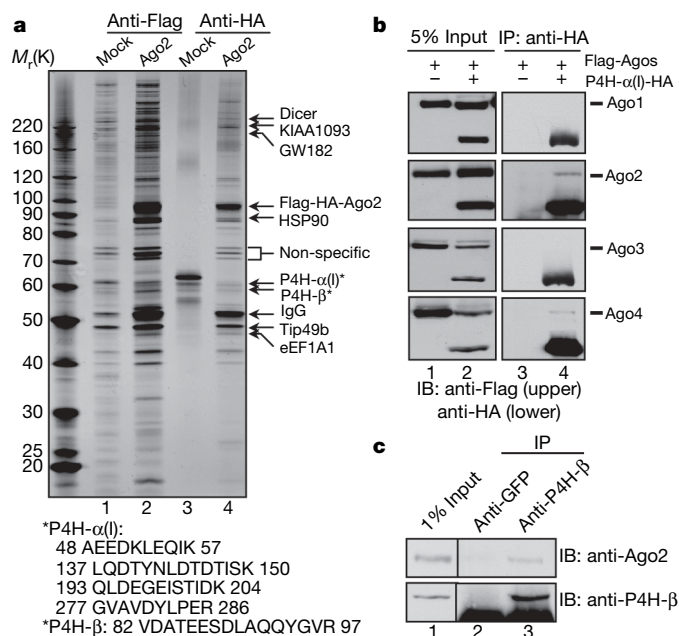


Figure 1 | Human Ago2 associates with C-P4H(I) subunits. **a**, Tandem affinity purification was conducted on cytoplasmic extracts from Mock (lanes 1 and 3) and Flag-HA-tagged Ago2 stable HeLa S3 cells (lanes 2 and 4) with sequential anti-Flag and HA immunoprecipitation. The peptides for P4H- α (I) and P4H- β are shown below. **b**, Flag-Ago1–4 were co-transfected with empty vector (lanes 1 and 3) or P4H- α (I)-HA (lanes 2 and 4) in 293ET cells. The immunoprecipitates with anti-HA antibody were blotted with anti-Flag and HA antibodies, respectively. IP, immunoprecipitate; IB, immunoblot. **c**, 293ET cells were immunoprecipitated with anti-GFP or P4H- β antibodies and followed with western blotting with anti-P4H- β and Ago2 antibodies.

¹Department of Pathology, Harvard Medical School, New Research Building 854, 77 Avenue Louis Pasteur, Boston, Massachusetts 02115, USA. ²Cutaneous Biology Research Center, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts 02129, USA. ³Oulu Centre for Cell-Matrix Research, Biocenter Oulu and Department of Medical Biochemistry and Molecular Biology, University of Oulu, FIN-90014 Oulu, Finland. ⁴Department of Human Genetics, Center for Neurodegenerative Disease, School of Medicine, Emory University, Atlanta, Georgia 30322, USA. ⁵Division of Endocrinology, Diabetes, and Hypertension, Department of Medicine and BCMP, Brigham and Women's Hospital and Harvard Medical School, 221 Longwood Avenue, Boston, Massachusetts 02115, USA.

in Ago2. They were conserved among other human Ago proteins (Supplementary Fig. 1) and Ago proteins from *Mus musculus*, *Drosophila melanogaster* (Ago1) and *Caenorhabditis elegans* (Agl1), but not *Arabidopsis thaliana* and *Schizosaccharomyces pombe*. Ago4 contains two additional X-P-G motifs at its amino terminus, of which the R-P-G (18–21) motif is conserved in Ago1 and 3. In addition, Ago2 has seven X-P-A triplets; five of them are conserved in human Ago1, 3, 4 and Ago proteins from *M. musculus* and *D. melanogaster*.

To determine the hydroxylation site(s) in Ago2, endogenous Ago2 was immunoprecipitated from 293ET cells and analysed by tandem mass spectrometry (MS/MS). The analysis identified one hydroxylation site at Pro 700 (DYQP*GITFIVVQK) in Ago2. The hydroxylated peptides (Fig. 2a, upper panel) represented approximately 20% (7,900/39,900) of total Ago2 peptides detected (DYQPGITFIVVQK). However, this is probably an underestimation because the MS/MS pattern of this hydroxylated peptide is very similar to that of its unmodified peptide (Fig. 2a, lower panel), as would have been expected for this type of modification. In contrast, we did not detect peptides containing modified proline at other X-P-G sites such as Pro 120 or 523, suggesting that Ago2 hydroxylation at Pro 700 is likely to be specific. This analysis, however, does not exclude cell type and tissue- and/or developmental-stage-specific hydroxylation at other proline residues of the Ago proteins.

Human Ago2 was originally reported as GERP95 (Golgi endoplasmic reticulum protein 95 kDa)⁵. Biochemical analysis suggests that Ago2 is a cytoplasmically exposed, peripheral membrane protein that exists in a protease-resistant complex⁵. Although Ago2 does not have either signal

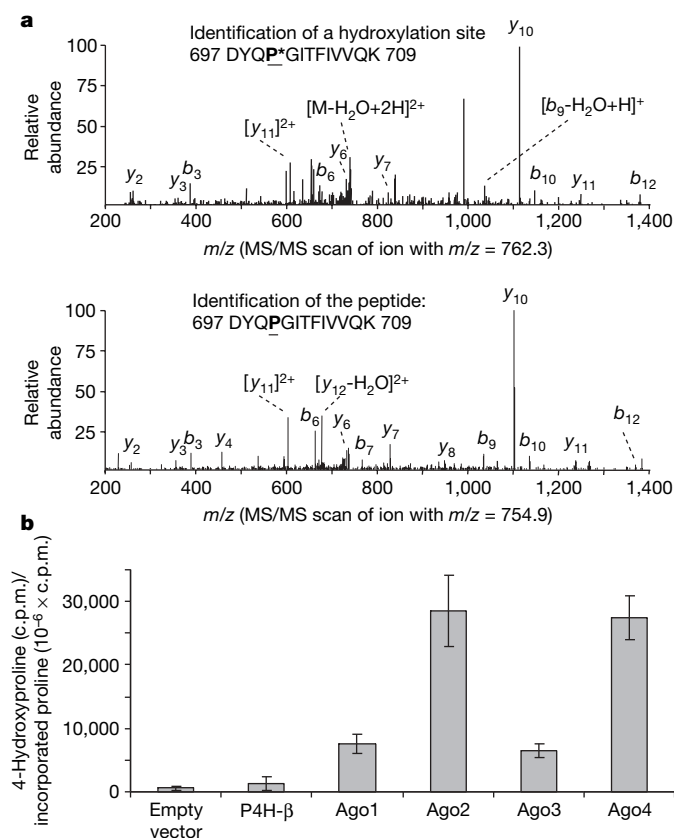


Figure 2 | In vivo and in vitro hydroxylation of Ago2. **a**, Endogenous Ago2 is hydroxylated at proline 700. Endogenous Ago2 was immunoprecipitated with a polyclonal anti-Ago2 antibody and analysed by LC-MS/MS. The MS/MS scanned hydroxylated peptides with ion m/z 762.3 (upper panel) and ion m/z 754.9 (lower panel). **b**, In vitro hydroxylation of Ago proteins. pcDNA3-Flag-HA-Ago1–4 were in vitro translated in the presence of L-[2,3,4,5-³H]proline. The in vitro translated lysates were incubated with purified recombinant C-P4H(I) enzyme and the reaction cofactors. The amount of 4-hydroxy[³H]proline formed in control, P4H-β and Ago 1–4 was determined as described in Methods. Data are presented as the mean \pm standard deviation for three independent experiments.

peptides or transmembrane domains, the expression of the exogenously introduced Ago2 overlaps with that of the ER marker BiP in NRK and COS cells⁵. Endogenous Ago2 is largely cytoplasmically localized, but is also enriched where P4H-β signal is found in a percentage of U2OS cells. Partial overlap with P4H-β was also observed in HeLa and 293ET cells (Supplementary Fig. 2). Importantly, the in vitro hydroxylation assay demonstrated that only C-P4H(I) but not the cytoplasmically localized prolyl-4-hydroxylases (PHD 1–3) was able to mediate Ago protein hydroxylation in vitro (Fig. 2b and data not shown). All four Ago proteins are hydroxylated in vitro by recombinant human C-P4H(I); however, Ago2 and Ago4 seem to be hydroxylated to a greater extent (Fig. 2b), consistent with the co-immunoprecipitation data that P4H-α(I) associated with Ago2 and 4 more readily than with Ago1 and Ago3 (Fig. 1b).

Hydroxylation of collagen and collagen-like proteins is essential for their functional triple-helical folding³. Inhibition of the C-P4H(I) enzyme generates unstable collagens, resulting in reduced collagen production⁶. In contrast, overexpression of P4H-α(I) is associated with excess collagen production⁷. To determine whether hydroxylation of proline at position 700 regulates Ago stability, we mutated the proline residue to alanine (P700A). As shown in Fig. 3a, Ago2 with the P700A mutation was less stable ($t_{1/2}$ approximately 6 h) than the wild-type protein whose half-life was estimated to be greater than 10 h (Fig. 3a). In contrast, P700A mutation had no effect on the stability of Ago1 and Ago3, suggesting that proline 700 is important for Ago 2 stability.

To determine whether the reduced stability of the Ago2 (P700A) mutant was due to loss of hydroxylation, we first compared the steady-state levels of the endogenous Ago2 in the wild type versus cells in which the expression of the hydroxylase, C-P4H(I), was inhibited by RNAi. We found knockdown of P4H-α(I) or -β reduced the steady-state level of the endogenous Ago2 protein in U2OS (Fig. 3b, lanes 4–7) and HeLa cells (data not shown) without affecting its mRNA level (Fig. 3b), suggesting that C-P4H(I) may regulate Ago2 protein stability. Upon MG132 treatment, the reduced endogenous Ago2 protein caused by P4H-α(I) RNAi was restored to the level comparable to that of the control RNAi sample (Fig. 3c), suggesting that reduced Ago2 is likely to be due to degradation by the proteasome-mediated pathway.

We examined the protein turnover rate of the endogenous Ago2 with or without an intact C-P4H(I). In the control RNAi U2OS cells, we found endogenous Ago2 to be quite stable with very little turnover, even at 10 h post-cycloheximide (50 $\mu\text{g ml}^{-1}$) treatment (Fig. 3d). In the cells where C-P4H(I) expression was inhibited by short hairpin RNA (shRNA), Ago2 became slightly destabilized at 4 h, and the steady-state level was further decreased at later time points (Fig. 3d). These results show that a reduction of the steady-state level of Ago2 is correlated with reduction of the hydroxylase C-P4H(I). Taken together, these findings suggest that hydroxylation of Ago2 at P700 regulates its stability.

The observation that Ago2 hydroxylation regulates its stability predicts that the hydroxylase C-P4H(I) is likely to play a role in RNAi. To address this issue, we assayed the requirements of C-P4H(I) for short interfering (si)RNA- or miRNA-programmed RISC (si or miRISC) activity. We first constructed and validated GFP (green fluorescent protein) reporters containing perfect (PE) or bulged (BU) complementary sequences for miRNA let-7 or miR21 (Supplementary Fig. 3). We next established U2OS stable cell lines expressing long GFP (GFPL) and GFP-let-7 (PE), GFPL and GFP-miR21 (8 × BU), respectively. The siRISC or miRISC activities were measured by determining the ratio between the reporter GFP-let-7 (PE) or GFP-miR21 (8 × BU) and the GFPL control. As shown in Fig. 4a, inhibition of Ago2, not Ago1, reduced the siRISC activity to 20% of that with the control RNAi. Knockdown of P4H-α(I) and -β reduced let-7 siRISC activity to approximately 40% (Fig. 4a). Reverse transcription-PCR (RT-PCR) of GFPL and GFP-let-7 (PE) showed derepression of the GFP-let-7 (PE) reporter at the RNA level (Fig. 4a). Consistently, P4H-α(I) null mouse embryonic fibroblast (MEF) cells⁸ also showed reduced Ago2 protein level and let-7-guided siRISC activity compared with wild-type MEF cells (Fig. 4c). In contrast, miRISC activity was not altered upon C-P4H(I) knockdown as measured by the GFP-miR21

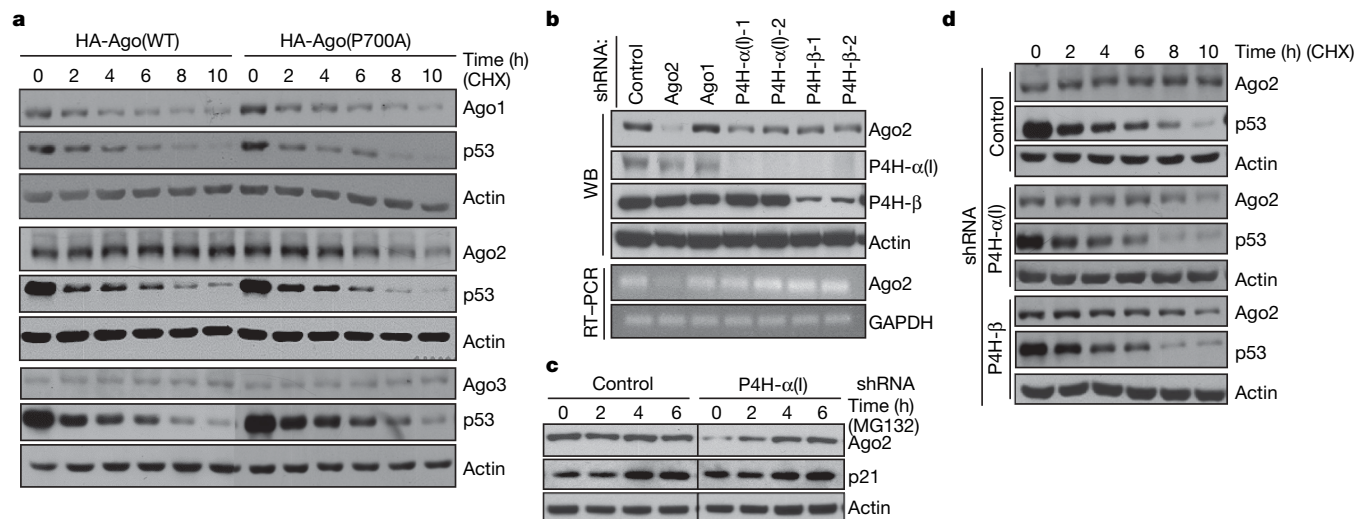


Figure 3 | Impaired hydroxylation downregulates Ago2 stability. **a**, U2OS cells transfected with HA-Ago1, 2, 3 (wild type, WT) and their mutants (P700A) were treated with 50 $\mu\text{g ml}^{-1}$ of cycloheximide (CHX) for the indicated times. Agos, p53 and actin were detected by western blot. The results shown are representative for three independent experiments. **b**, U2OS cells were transfected with indicated shRNAs followed by puromycin selection. Western blot with indicated antibodies and RT-PCR were

performed. **c**, P4H- α (I)-depleted and control U2OS cells were treated with 25 μM of MG132 for the indicated times. Endogenous Ago2, p21 and actin were detected by western blot. **d**, U2OS cells were transfected with P4H- α or - β shRNA followed with puromycin selection. Cells were then treated as in **a**. Ago2, p53 and actin were monitored by western blot. The results are representative of three independent experiments.

(8 \times BU) reporter (Fig. 4b), although whether this observation is generally applicable to other miRNA-mediated translational inhibition remains to be determined. Hydroxylation appears to be dispensable for the catalytic activity of Ago2 as the Ago2 hydroxylation mutants were capable of restoring siRISC activity upon overexpression in cells in which the endogenous Ago2 expression was inhibited by an shRNA targeting the 3' untranslated region of Ago2 (Fig. 4d). Furthermore, bacterially produced Ago2 binds single-stranded siRNAs and is able to cleave siRNA targets, suggesting that hydroxylation is not required for

Ago2 siRNA binding and slicing activity because bacteria do not have C-P4H (ref. 9). Taken together, C-P4H(I) appears to play an important role in RNAi by regulating Ago2 stability.

It has been reported that Ago2 is localized in cytoplasmic mRNA processing bodies (P-bodies)¹ and is recruited to stress granules in response to oxidative stress or specific translational inhibitors¹⁰. Although the P-body does not appear to be required for RNAi¹¹, it is important for RNA-related functions in controlling mRNA translation and degradation¹². Thus, Ago2 P-body localization may play a

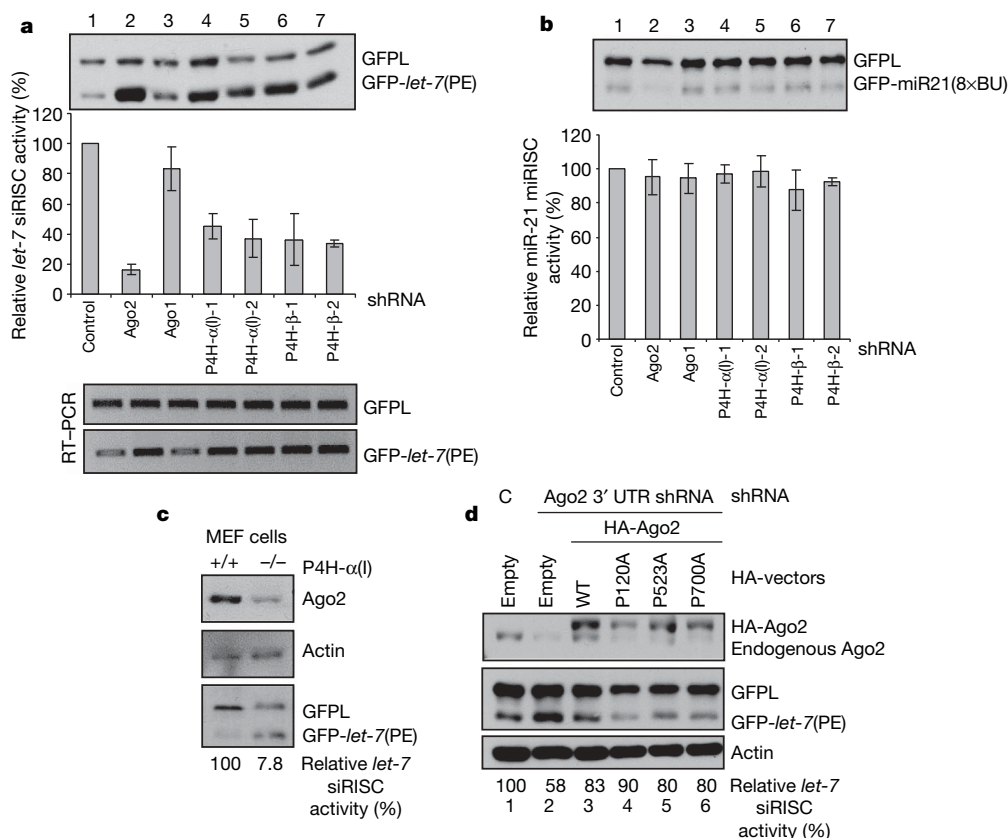


Figure 4 | Impaired C-P4H(I) reduced *let-7* guided siRISC activity. **a**, **b**, U2OS stable cell line expressing GFP/let-7 (PE) (**a**) or GFP/let-7 (PE) (8 \times BU) (**b**) were transfected with indicated shRNAs and selected by puromycin. The expression ratios between GFP and GFP reporter were calculated and the intact (100%) si/miRISC activity was considered in control shRNA cells. The relative si/miRISC activities under knockdown conditions of the indicated genes were obtained by comparison with the control shRNA transfection. The standard deviations were obtained from three independent experiments. **c**, Control and P4H- α (I) null MEF cells were transfected with GFP and GFP-let-7 (PE) plasmids. Ago2, actin and GFP expressions were detected by western blot. **d**, Ago2(P700A) is catalytically active. GFP/let-7 (PE) stable cell line was co-transfected with control shRNA (lane 1) or shRNA targeting the 3' untranslated region of Ago2 (lanes 2–6) together with empty vector (lanes 1 and 2) or HA-Ago2 constructs as indicated. GFP, Ago2 proteins were detected. *let-7* guided siRISC activity was calculated as in **a**.

yet-to-be-identified function in RNA-mediated gene regulation. Therefore, we investigated whether hydroxylation has any impact on Ago2 P-body and stress-granule localization. We found that depletion of the C-P4H(I) subunits by RNAi or genetic ablation of P4H- α (I) in MEF cells reduced Ago2 P-body localization marked by Dcp1a, which resides in P-bodies^{12–14} (Supplementary Fig. 4a). The reduced P-body localization of Ago2 might be a reflection of the reduced steady-state level of Ago2 protein when the expression of the hydroxylase is impaired. We also found that P700A mutation caused the mutant proteins to localize away from Dcp1a, compared with wild type and the other proline mutants of Ago2 (P120A and P523A) (Supplementary Fig. 4b). P700A also delocalized Ago1 from P-bodies. In contrast, P700A did not cause delocalization of Ago3 from P-bodies, but reduced Ago4 P-body localization (Supplementary Fig. 4b), suggesting that proline 700 may impact different Ago proteins differently. Affinity purification and co-immunoprecipitation from wild-type and mutant (P700A) Ago2 showed unchanged Dicer but reduced Dcp1a and KIAA1093 association with Ago2 (P700A) (Supplementary Fig. 4c, d). This finding suggests that Ago1 and 2 may be regulated similarly in their P-body localization. To ask whether hydroxylation has any effects on Ago2 stress-granule re-localization, we treated U2OS cells containing either intact or impaired C-P4H(I) with arsenite (Supplementary Fig. 5), translation inhibitor (hippuristanol) or heat shock (data not shown) and found no dramatic changes in Ago2 stress-granule re-localization.

Hydroxylation is an important post-translational modification, and plays important biological roles. C-P4H-mediated hydroxylation of collagen proteins is essential for their functional folding and stability⁵. Proline hydroxylation of the hypoxia-inducible factor (HIF) α -subunit by a separate P4H family causes recruitment of the von Hippel Lindau (VHL) ubiquitin ligase complex for HIF- α degradation¹⁵. We have shown that hydroxylation regulates Ago2 stability, suggesting a novel role for C-P4H. It will be important to investigate whether this modification is conserved in other species. For example, proline 700 is conserved in *C. elegans* *Alg1*, *Alg2* and T23D8.7 proteins, which function in miRNA biosynthesis¹⁶. Additionally, the P4H- α subunit is rate-limiting in the formation of active C-P4H. Several pathways or stimuli regulate P4H- α (I) at transcriptional and post-transcriptional levels. For example, transforming growth factor (TGF)- β 1 stimulates transcription of the P4H- α (I) gene¹⁷. Hypoxia upregulates the mRNA levels of P4H- α (I) and (II)¹⁸. At post-transcriptional level, the RNA-binding protein nucleolin has been demonstrated to bind to the 5' and 3' untranslated regions of P4H- α (I) mRNA and enhance P4H- α (I) translation efficiency under hypoxia¹⁹. Further efforts to delineate if the P4H- α regulators might also regulate RNAi machinery through regulation of hydroxylation of Ago proteins will provide additional insights into RNAi regulatory mechanisms.

METHODS SUMMARY

Tandem affinity purification, *in vivo* hydroxylation and mass spectrometry. The tandem affinity purification was performed as described previously²⁰. Briefly, HeLa S3 stable cell lines expressing Flag–HA-tagged Ago proteins were cultured in 32 litres. Cytoplasmic extracts were immunoprecipitated with an anti-Flag M2 antibody (Sigma) followed by anti-HA antibody resins (Santa Cruz) with buffer A (20 mM Tris, pH 8.0, 100 mM KCl, 5 mM MgCl₂, 0.2 mM EDTA, 10% glycerol, 0.1% Tween-20, 10 mM 2-mercaptoethanol and 0.25 mM phenylmethylsulphonyl fluoride (PSMF)). HA immunoprecipitation elution 1 (50 μ l) was analysed by mass spectrometry using an LTQ-FT hybrid mass spectrometer (Thermo Finnigan). To identify the hydroxylation site(s) of endogenous Ago2, immunoprecipitation was performed using an anti-Ago2 polyclonal antibody (Upstate) from 293ET cells (see immunoprecipitation condition and buffer in Methods). The Ago2-containing band was cut from SDS–polyacrylamide gel electrophoresis gel, fully trypsinized and analysed by reverse-phase liquid chromatography–tandem mass spectrometry (LC–MS/MS) as previously reported²¹. An MS/MS scan of the precursor ion *m/z* 762.3, in which it was fragmented into multiple labelled product ions (*b* and *y* ions), identified a peptide according to the mass shift (+16 Da) caused by hydroxylation. An MS/MS scan of the precursor ion *m/z* 754.9 indicated the peptide without modification.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 April; accepted 20 June 2008.

Published online 6 August 2008.

- Peters, L. & Meister, G. Argonaute proteins: mediators of RNA silencing. *Mol. Cell* **26**, 611–623 (2007).
- Robb, G. B. & Rana, T. M. RNA helicase A interacts with RISC in human cells and functions in RISC loading. *Mol. Cell* **26**, 523–537 (2007).
- Myllyharju, J. & Kivirikko, K. I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet.* **20**, 33–43 (2004).
- Kivirikko, K. I., Myllylä, R. & Pihlajaniemi, T. in *Post Translational Modification of Proteins* (eds Harding, J. J. & Crabbe, M. J. C.) 1–51 (CRC Press, 1992).
- Cikaluk, D. E. *et al.* GERp95, a membrane-associated protein that belongs to a family of proteins involved in stem cell differentiation. *Mol. Biol. Cell* **10**, 3357–3372 (1999).
- Rocnik, E. F., Chan, B. M. & Pickering, J. G. Evidence for a role of collagen synthesis in arterial smooth muscle cell migration. *J. Clin. Invest.* **101**, 1889–1898 (1998).
- John, D. C. *et al.* Expression of an engineered form of recombinant procollagen in mouse milk. *Nature Biotechnol.* **17**, 385–389 (1999).
- Holster, T. *et al.* Loss of assembly of the main basement membrane collagen, type IV, but not fibril-forming collagens and embryonic death in collagen prolyl 4-hydroxylase I null mice. *J. Biol. Chem.* **282**, 2512–2519 (2007).
- Rivas, F. V. *et al.* Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Struct. Mol. Biol.* **12**, 340–349 (2005).
- Leung, A. K. L. C. J. & Sharp, P. A. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl Acad. Sci. USA* **103**, 18125–18130 (2006).
- Chu, C. Y. & Rana, T. M. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol.* **4**, e210 (2006).
- Parker, R. & Sheth, U. P bodies and the control of mRNA translation and degradation. *Mol. Cell* **25**, 635–646 (2007).
- Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nature Cell Biol.* **7**, 719–723 (2005).
- Sen, G. L. & Blau, H. M. Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nature Cell Biol.* **7**, 633–636 (2005).
- Kaelin, W. G. Jr. The von Hippel-Lindau protein, HIF hydroxylation, and oxygen sensing. *Biochem. Biophys. Res. Commun.* **338**, 627–638 (2005).
- Grishok, A. *et al.* Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**, 23–34 (2001).
- Chen, L. *et al.* Human prolyl-4-hydroxylase α (I) transcription is mediated by upstream stimulatory factors. *J. Biol. Chem.* **281**, 10849–10855 (2006).
- Hofbauer, K. H. *et al.* Oxygen tension regulates the expression of a group of procollagen hydroxylases. *Eur. J. Biochem.* **270**, 4515–4522 (2003).
- Fahling, M. *et al.* Translational control of collagen prolyl 4-hydroxylase- α (I) gene expression under hypoxia. *J. Biol. Chem.* **281**, 26089–26101 (2006).
- Nakatani, Y. & Ogryzko, V. Immunoaffinity purification of mammalian protein complexes. *Methods Enzymol.* **370**, 430–444 (2003).
- Peng, J. & Gygi, S. P. Proteomics: the move to mixtures. *J. Mass Spectrom.* **36**, 1083–1091 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank C. C. Mello, A. Grishok and P. A. Sharp for reading this manuscript and discussion. We thank E. Lehtimäki and R. Juntunen for *in vitro* hydroxylation assays, J. Lykke-Andersen and N. Kedersha for polyclonal anti-Dcp1a antibody and anti-Tial antibody, and the personnel of the Biocenter Oulu Transgenic Animal Core Facility and the University of Oulu Laboratory Animal Centre for technical assistance. We also thank the Kazusa DNA Research Institute (Japan) for providing the KIAA clones with numbers 1567, 1093, 1582 and 1460 and N. R. Wall for pOZ-FH-Ago1 plasmid. H.H.Q. was supported by a Canadian Institute of Health Research postdoctoral fellowship. This work was supported by grants from the Health Science Council of the Academy of Finland (202469) and the S. Jusélius foundation to J.M. and National Institutes of Health grants to J.P. (AG025688) and Y.S. (GM53874).

Author Contributions H.H.Q. performed tandem affinity purification, Co-IP, Ago2 stability, si/miRISC activity, P-body and stress-granule experiments; P.P.O. established GFP reporter stable cell lines and performed MG132 experiments in the Lee laboratory; D.C. and J.P. performed mass-spectrometric analysis for the tandem affinity purified Ago complexes and hydroxylation; J.M. and O.P. conducted *in vitro* hydroxylation and provided P4H- α (I) null MEF cells; Yujiang Shi helped in tandem affinity purification. H.H.Q. and Yang Shi conceived and designed the study, and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.S. (yang_shi@hms.harvard.edu).

METHODS

Plasmid constructions, antibodies and chemicals. Human Ago4 (KIAA1567), KIAA1093, KIAA1582 and GW182 (KIAA1460) cDNAs were obtained from the Kazusa DNA Research Institute (Japan), the others were amplified by RT-PCR. For tandem affinity purification, pOZ-FH-N-Ago1–4 were constructed by inserting the open reading frames into the *XhoI/NotI* S sites of pOZ-Flag-HA-N vector. For pEGFP1-Agos, *Sall/BamHI* and *Sall/SmaI* were chosen for Ago1, 2, 3 and 4 respectively. To generate amino-terminal Flag-tagged Agos, Ago1, 2, 3 and 4 were cloned into *HindIII/NotI* and *NotI/Sall* sites of 3 × Flag-CMV-7.1. To express Flag-HA-tagged Agos in pcDNA3 vector, the coding regions and Flag-HA (FH) tags were PCR-amplified from pOZ-FH-N-Agos and cloned into *EcoRI/NotI* (Ago1, 2 and 3) and *KpnI/NotI* (Ago4) sites of pcDNA3 vector. For HA-tagged vectors, KIAA1093 was cloned into *BamHI/XbaI* sites of pcDNA-C-HA. Point mutations were introduced following standard mutagenesis procedures. Polyclonal anti-P4H- α (I) was generated (Innovagen) against purified recombinant P4H- α (I) peptide-substrate-binding domain²². Polyclonal anti-P4H- β antibody, polyclonal anti-TIA-1, monoclonal anti-CDC25a, monoclonal anti-GFP(B2) and monoclonal anti-p53 antibodies were purchased from Santa Cruz. Rabbit polyclonal anti-Ago2 antibody was obtained from Upstate, and polyclonal anti-Dcp1a antibody was a gift from J. Lykke-Andersen. Monoclonal antibodies for Ago2 (Novus Biochemicals and Wako Chemicals), Dicer (Abcam), p21 (Calbiochem), HA (Covance) and Flag M2 (Sigma) were also used. Cycloheximide (Sigma), Z-Leu-Leu-Leu-al (MG132) (Sigma), arsenite (Riedel de Haen) and hippuristanol (a gift from the P. Sharp laboratory) were obtained.

In vitro hydroxylation. pcDNA3-Flag-HA-Ago1–4 were translated *in vitro* in the presence of radioactively labelled proline, L-[2,3,4,5-³H]proline (85 mCi mmol⁻¹) in rabbit reticulocyte lysate using TNT® Quick-coupled transcription/translation system (Promega). The empty vector and pcDNA-P4H- β were used as negative controls. The *in vitro* translated lysates were aliquoted into two portions for C-P4H(I) hydroxylation reaction in which 0.33 pmol of recombinant C-P4H(I) purified from insect cells²³ or purified HIF PHD (1–3)²⁴ was added. The enzyme was omitted from the parallel tubes that served as negative controls. The reactions were dialysed extensively to remove any remaining free radioactive proline. The amount of 4-hydroxy[³H]proline formed in the substrate was determined by a specific radiochemical procedure as previously described²⁵.

Endogenous miRNA-guided si/miRISC activity assay, shRNAs and RT-PCR. To generate the reporter system for miRNA-guided siRISC and miRISC activity reading, pcDNA3-GFP was first cloned by inserting the GFP coding region into *BamHI/NotI* sites of pcDNA3. Subsequently, one copy of perfect complementary sequence of miRNA *let-7* (5'-AACTATACAACCTACTACCTCA-3') or eight copies of bulged complementary sequence of miR21 (5'-TCAACATCAGA GAGATAAGCTA-3') were cloned into the *NotI/XhoI* sites of pcDNA3-GFP to form pcDNA3-GFP-*let-7* (PE) and pcDNA3-GFP-miR-21 (8 × BU). pcDNA3-GFPL was obtained by removing the stop codon from pcDNA3-GFP and produces a mutant long GFP (GFPL) (267 amino acids compared with its original form of 239 amino acids) serving as a control. pcDNA3-GFPL was co-transfected with pcDNA3-GFP-*let-7* (PE) or pcDNA3-GFP-miR-21 (8 × BU) into U2OS cells followed by 2 mg ml⁻¹ G418 selection for two weeks. The resistant colonies were propagated into a stable cell line. shRNAs used in this study were all cloned into *XhoI/BamHI* sites of pBabe-U6 vector, which also

expresses a puromycin-resistant gene²⁶. The sequences of shRNAs (passenger strand, 5' to 3') were as follows: Ago1 (GATGAAGAATGCCAGCTACAA), Ago2 (ORF) (GGCACAGCCAGTAATCGAGTT), Ago2 (3' untranslated region) (GCTACACTCAGACCAACAGAT), P4H- α (I)-1 (GTATTATTCGCT TCCATGATA), P4H- α (I)-2 (GGCTAACTAGTACAGCGACAA), P4H- β -1 (GCCGACAGGACGGTCATTGAT), P4H- β -2 (GATGAAGTGAATACGCA AAG) and control (GAACGTCATCAAGCTGATCTA). For si/miRISC activity assay, 2 μ g of corresponding shRNAs were transfected into 5 × 10⁵ cells of the GFP-reporter stable cell lines in six-well plates, the cells were transferred to 10-cm plate after 24 h and selected with 2 μ g ml⁻¹ of puromycin for 36 h. For RT-PCR, total RNAs were extracted using Trizol reagents (Invitrogen) from either GFP reporter cell lines, or U2OS cells transfected with corresponding shRNAs. Primers used for RT-PCR as follows: common forward primer for GFPL and GFP-*let-7* (PE): GAACGGCATCAAGGTGAACCTT, specific reverse primers for GFPL: TAGCGTAATCTGGACATCGTATGGGT, and GFP-*let-7* (PE): GACGACCTCGAGTGAGGTAGTAGTTGTATA, GAPDH: forward GAAGG TGAAGGTCGGAGTC and reverse GAAGATGGTGATGGGATTTC, Ago2: forward GATCGCAAGAAGAGATTA and reverse CGTCTGCCGCCAGG ATGAC.

Immunoprecipitation and immunofluorescence. Three × Flag-Ago constructs and HA-P4H- α (I) as well as 3 × Flag-wtAgo2, Ago2(P700A) and KIAA1093-HA were co-transfected into 293ET cells. Cells were lysed 48 h after transfection in lysis buffer containing 300 mM NaCl, 50 mM TrisHCl, 5 mM EDTA, 0.1% NP-40, 0.5 mM NaF, 0.5 mM Na₃VO₄ and 1 × proteinase inhibitor (Roche). The lysates were incubated with anti-HA resins (Santa Cruz) for 3 h and washed five times with washing buffer (lysis buffer with 150 mM NaCl). The immunoprecipitated products were separated by SDS-polyacrylamide gel electrophoresis and blotted with anti-HA and anti-Flag antibodies, respectively. To assay P-body localization of Ago2, U2OS cells were first transfected with shP4H- α (I)-2 and - β -2 followed by puromycin selection for 36 h. The selected cells were seeded on coverslips for immunofluorescence. Subsequently, the cells were fixed with 3.7% paraformaldehyde in PBS for 30 min followed by permeabilization by 0.2% Triton X-100. The samples were then incubated with either anti-HA or Ago2 monoclonal antibodies at 1:500 dilutions simultaneously with anti-Dcp1a polyclonal antibody at 1:1,000 dilutions. Secondary antibodies against mouse immunoglobulin G conjugated with Alexa Fluor 488, rabbit IgG with Alexa 594 and Hoechst 33258 (Molecular Probes) were applied correspondingly after three washes with 0.1% Tween-20 in PBS.

22. Hieta, R. *et al.* The peptide-substrate-binding domain of human collagen prolyl 4-hydroxylases. Backbone assignments, secondary structure, and binding of proline-rich peptides. *J. Biol. Chem.* **278**, 34966–34974 (2003).
23. Vuori, K. *et al.* Characterization of the human prolyl 4-hydroxylase tetramer and its multifunctional protein disulfide-isomerase subunit synthesized in a baculovirus expression system. *Proc. Natl. Acad. Sci. USA* **89**, 7467–7470 (1992).
24. Hirsilä, M. *et al.* Effect of desferrioxamine and metals on the hydroxylases in the oxygen sensing pathway. *FASEB J.* **19**, 1308–1310 (2005).
25. Berg, R. A. Determination of 3- and 4-hydroxyproline. *Methods Enzymol.* **82**, 372–398 (1982).
26. Liu, G., Xia, T. & Chen, X. The activation domains, the proline-rich domain, and the C-terminal basic domain in p53 are necessary for acetylation of histones on the proximal p21 promoter and interaction with p300/CREB-binding protein. *J. Biol. Chem.* **278**, 17557–17565 (2003).

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

Why do postdocs in many countries have such a hard time working to improve their lot? The answer lies, in part, with the nature of the postdoctoral experience. Unlike graduate students, who are increasingly forming unions (see *Nature* **451**, 861; 2008), the two-year duration of most postdoc positions is often not enough time to seriously engage with issues such as salary, stipends and working conditions. It's hard for postdocs to invest time in a venture that will gain them few benefits, and will distract from their main objectives: conducting and publishing research and landing a job in what continues to be a very competitive landscape. Two years is also little time to drill through the dense red tape of universities and governments.

Yet, as detailed in an article on page 426, some postdoctoral associations are showing signs of progress, however modest. Perhaps the greatest success story has been the US National Postdoctoral Association (NPA). Started five years ago, the NPA has shown what can be done given ample outreach and good seed funding from prominent backers such as the Alfred P. Sloan Foundation in New York. At the NPA's annual meeting held in April in Boston, Massachusetts, the Sloan Foundation's Michael Teitelbaum recalled how he was initially not sure whether the association would make it, given the challenges of attracting a dedicated membership to an untested model. He has been pleasantly surprised, he says, at how much it has grown and the degree to which it is now self-sustaining. Walter Schaffer, a senior scientific adviser for extramural research at the National Institutes of Health (NIH), says that the NPA hasn't yet generated major changes in policy, but it has garnered a voice that makes postdocs feel less disenfranchised.

The NPA might have the ear of important institutions such as the NIH, but like many unions, it still faces frustrating bureaucratic roadblocks. Nevertheless, the NPA, and hence the US postdoc, has a voice that should remain strong and consistent for the foreseeable future. It is to be hoped that other nations, and other nations' postdocs, can follow its example.

Gene Russo is editor of *Naturejobs*.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com

Natureevents:
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com

UK Corporate:
Nils Moeller (4953)

Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Reya Silao (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock
Naturejobs online production: Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718

Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com



Strength in bonding

It is not in numbers, but in unity, that our great strength lies,” wrote Thomas Paine in his pamphlet *Common Sense*. His words, published in 1776, galvanized the American colonies in the fight for independence from Britain. That same sentiment applies to the growth of postdoc organizations in recent years. But despite the rise in numbers of postdoc associations (PDAs) and membership, they are still struggling to effect major policy change. Many PDAs — whose members seek better pay, benefits and career training — are making some headway locally (see ‘Showing initiative’), but change on a regional level is a far bigger challenge.

The birth of the National Postdoctoral Association (NPA) in the United States helped increase the number of PDAs at universities by 60–70%. In Europe, Eurodoc, which addresses both graduate student and postdoc issues, has grown to include 44 countries as either members, observers or networkers. Meanwhile, universities and research institutions across the United States, Canada, Europe and Japan are working to add associations every year.

Still, the useful change resulting from these efforts

Postdocs are attempting to organize and improve their lot worldwide — but results have been mixed. Paul Smaglik reports.

is minimal. The recent history of the NPA illustrates what such associations can accomplish — and how their power is limited. The most gratifying part of running the NPA for five years, former executive director Alyson Reed says, was bringing together and inspiring fellows who, before they had their own association, felt voiceless. “We created a community where before there was a lot of isolation and lack of connections,” Reed says.

Reed says that the NPA has helped universities establish their own postdoc offices (PDOs), which connect to regional associations. The PDA–PDO links help address local issues, such as the need for better connections to business and career opportunities in local industry, and campus issues, such as access to childcare.

But Reed laments the NPA’s failure to help increase stipend levels for US postdocs. The NPA wrote a policy paper, testified in Congress and had representatives meet with Elias Zerhouni, director of the US National Institutes of Health (NIH) — all to no avail. The NIH is one of the biggest funders of postdocs in the United States and helps set the standard for stipends in government-sponsored research. “People made



promises — people in leadership positions and government agencies. They are just falling far short of those pledges,” Reed says. “Postdocs have been very patient and very collegial and trying to do things the polite way. But there comes a time when you have to say ‘words are nice, but they don’t pay the bills.’”

Between 2000 and 2004, the agency raised entry-level postdoc stipends from \$26,900 to \$36,900. There’s been no significant increase since then, but the NIH has improved postdoc healthcare and childcare benefits, notes Norka Ruiz Bravo, NIH deputy director for extramural research.

American model

The NPA’s influence has already spread beyond the United States to Europe. Postdocs see it as a model, and some even use the ‘tool kit’ the association provides to help fellows start up their own PDAs. John Bothwell, a postdoc at the Marine Biological Association in Plymouth, England, has contacted and visited members of the NPA governing board hoping to form a UK postdoc association. Results have been mixed. So far, Bothwell has about 30 universities on board — a fraction of the United Kingdom’s total. Funds are a problem, although the UK government has shown some willingness to help. Included among the 2002 Roberts Review — a series of recommendations to foster science, engineering and innovation

SHOWING INITIATIVE

Although postdoc organizations worldwide have struggled to change national policy, they have been more successful highlighting issues specific to their individual constituencies. Postdoc organizations may have a hard time improving, for example, short-term contracts, benefits and stipends. But they have seen some success bringing in career-training programmes. Here is a sample of their ongoing initiatives.

Second that mentor

This autumn, the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, is encouraging fellows to choose another mentor beyond the principal investigator they work under day-to-day. “This gives the fellow some outside input when a postdoc and lab head don’t agree on the project or understand each other,” says Detlev Arendt, EMBL’s postdoc coordinator. Adding another adviser need not be burdensome for the faculty of the fellow: EMBL recommends that the two meet once in the first three months of the fellowship, then once or twice a year after that.

Survey says

The Young European Biotech Network (YEBN) this autumn is completing a survey of 1,000 members and 500 industry contacts about the skills needed in biotechnology and what young scientists can do to prepare for a biotech career. “There are gaps in the education and what industry expects,” says Francesco Lescai, YEBN chairman and a medical biotechnologist at the University of Bologna, Italy. He hopes the survey will help young scientists fill those gaps. Lescai also says that European biotech regulations vary by country. “There are certain professions in some countries that aren’t regulated at all, and in other countries they are very regulated.” The survey will help clarify what those differences are and make scientists more aware of them, Lescai says. The YEBN is also planning a meeting in early 2009 to conduct focus groups with members of YEBN, the European Federation of Biotechnology and the European Commission (EC) to determine the hot issues. This will result in a white paper to the EC about how to better prepare young scientists for biotech careers.

Taxation representation

The Canadian government, eager to import new science talent and keep local

stars home, has bolstered its scholarship schemes in recent years — offering as much as Can\$50,000 a year for some young students. Although Stuart Netherton, chair of the postdoctoral association at the University of Calgary, applauds the generous funding, he suggests that this represents a double bind for postdocs. First, some Canadian fellows earn less than graduate students receiving the new scholarships, says Netherton. Second, stipends are taxed but scholarships are not — meaning graduate students could theoretically bank about twice the money of their more senior scientists-in-training. The Calgary postdoc association (PDA) is lobbying the government about this gap and hopes to get other Canadian PDAs on board.

Fellows receiving EC funding face similar problems in their home countries. Different nations treat their tax status differently, says Yegor Domanov, secretary-general of the Marie Curie Fellows Association. Domanov says this is the biggest complaint he hears from Marie Curie fellows.

Career scheming

Only a small percentage of postdocs will land academic jobs. At the College of Life Sciences at the University of Dundee, Scotland, that number is 6–7%, says Nicola Stanley-Wall, who runs the postdoc office. So when the university set up that office about five years ago, to coordinate with its PDA, the office decided to focus on jobs outside academia, such as work with contract research organizations that run clinical and preclinical trials. “Our postdoc career development scheme helps highlight different areas of career development that weren’t necessarily

highlighted before,” Stanley-Wall says.

EMBL’s postdoc association offers similar services, focusing on general career development, says Arendt. “There are language courses.

There’s a workshop about preparing for the job market,” he explains. “Before they apply, they give a fake seminar and they are the fake committee. We give feedback on their written presentation and their oral presentation.”

Yasuo Kanematsu, a biology professor at Osaka University, Japan, is pushing for a government-backed programme that more closely integrates training for both academia and industry. “If we separate the two paths that’s not good,” says Kanematsu, who has had difficulty getting postdocs to organize.

P.S.

“There are gaps in education and what industry expects.”
— Francesco Lescai

— were suggestions to improve the science and technology pipeline. One called for better retention of postgraduate scientists. As a result, UK policy-makers authorized £6 million (US\$10.6 million) over five years to help improve postdoc conditions.

However, most UK universities averaged a few thousand pounds apiece, with little direction in how to spend the money, Bothwell says. It's mainly the larger institutions — such as the universities of Cambridge, Oxford, Manchester and Dundee — that have active, well-organized PDAs; smaller ones have fewer active associations. Bothwell says his next step is convincing participants from 30 or so UK PDAs to meet in December. He wants to know their goals and arrive at a consensus on how to achieve them.

Dealing with difference

Achieving consensus in Europe is complicated by the fact that funding schemes from national governments, the European Union (EU) and foundations differ. There is also a wide discrepancy among more developed systems in the north, such as those in France, Germany and Switzerland, and the south, such as in Spain and Italy. Part of the problem lies with the EU's inability to enforce its own policies, according to Paulo Silva, now a postdoc in biochemistry and molecular biology at Pennsylvania State University. Silva helped launch Eurodoc in 2004, but is frustrated by its failure to create widespread change. He notes one EU provision, launched in 2005, about research conduct that attempted to put graduate students, postdocs and faculty members on a more even playing field with regard to contracts and benefits. "Everybody should be treated as professionals, no matter what the stage," Silva says of the provision. But, he says, so far there's been no action.

Sometimes member nations clash when EU initiatives are misinterpreted — and the postdoc can be the one who suffers. Silva points to the Bologna Process, which aims to make PhD programmes more uniform in terms of length and quality across Europe, as one example. The EU recommends shortening the length of PhD programmes, yet some countries have interpreted that as shortening the amount of time they will support students and postdocs.

Many fellows active in European PDAs complain about a lack of oversight. EU policies cover only people with EU funds; fellows paid for by national schemes are subject to different regulations. But Detlev Arendt, head of the postdoc office at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, says that



Paulo Silva, who helped set up Eurodoc, is frustrated by its failure to effect wider change.

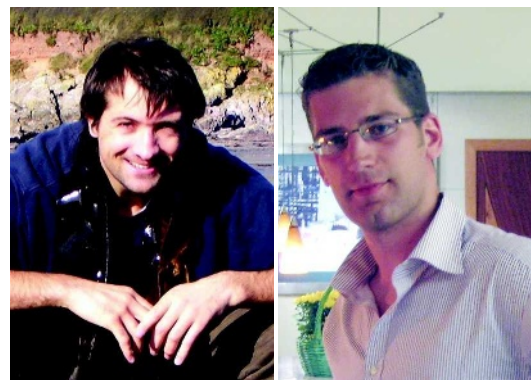
as the commission now funds about 40% of all postdocs in Europe, its influence is growing. And Arendt believes that even the European Charter for Researchers, which Silva criticizes as being toothless, has some merit. EMBL endorsed the charter, which sets rules of conduct for researchers. By doing so, EMBL improves its chances of earning EU funding — and attracting fellows who know they will be treated fairly under the commission's rules. "They set clear standards for working conditions," Arendt says.

But despite such charters, getting local institutions in line with EU policies is difficult, according to Yegor Domanov, secretary-general of the Marie Curie Fellows Association, and a postdoc at the University of Helsinki. Taxation is the biggest complaint, says Domanov — especially regarding supplementary income, such as mobility allowances and travel expenses, which are taxed by some host countries as income, even though the EU stipulates they should be considered expense reimbursement, not pay. Taxation of postdocs is also an emerging problem in Canada (see 'Showing initiative').

Domanov is collecting what he calls "horror stories" of postdoc experience, as well as other problems, such as visa issues for spouses, finding housing and language courses for foreign fellows. He aims to provide these to the EU in the hope that it will change its policies — or, at least, enforce existing ones better. Domanov echoes many fellows who say that creating real change by postdocs for postdocs is difficult — especially because of the transient nature of their fellowships versus the length of time it takes to effect real change. When fellows active in PDAs move on, many of the issues they have worked to address remain unresolved. Or else they get caught up in writing up their research, pursuing funding or landing their first permanent position — leaving them to pass the torch to the next generation.

Still, Domanov sees the rise of more European postdoc associations and their overall growth in membership as a sign that postdocs can make things better for themselves. "Little by little, people learn they have the power to change things," Domanov says.

Paul Smaglik is a freelance writer based in Milwaukee, Wisconsin.



John Bothwell (left) plans a UK-wide postdoc association; Yegor Domanov, from the Marie Curie Fellows Association.

The chess players

An unusual move.

Dan Gollub

Advances in biomedical technology had not been in the forefront of Paul's consciousness when he entered the spacious living room of Stacy Ashland's home. He'd been wondering if he might find a girlfriend at her party. The odds, he knew, were against it. He was struggling to earn a living as a chess player, and the other guests at her party were likely to have a higher status in society. And anyway, his introverted personality and an overall lack of experience in relationships tended to thwart his chances of romance. But all such thoughts vanished from his mind when he saw a woman present who'd been much in the news lately. Rather, she and Jarvis McKnight together had been featured in newscasts, interviews and other media presentations. In the centre of the room, surrounded by a dozen or so friends and admirers, was Lena Adrian, with Jarvis's head transplanted onto her right shoulder. "Phenomenal breakthrough," a popular newscaster had said.

Stacy was amid the partygoers surrounding Lena and Jarvis, and she saw Paul and waved at him to come over. "Jarvis," she said in her extravagant socialite's voice, "this is the chess player I mentioned to you. He's been teaching chess to my son the past three months."

Jarvis's head nodded sombly. "Chess has become more important to me recently."

He had a deep voice, Paul noted. The transplantation hadn't changed that aspect of his physiology.

"I'm sure the two of you will get along fine," Stacy commanded. She joined a nearby group of people and, as the hostess, was instantly welcomed into the conversation.

"Chess has an addictive side to it," Paul said. He saw that Lena was talking with a slim, long-legged man with a tattoo of a soccer ball on his arm. "Sometimes that's good, I suppose. And sometimes I lie awake after chess tournaments thinking about the moves I should have played differently."

Jarvis chuckled. Then a shadow crossed over his face. "Lena's not likely to sit through chess tournaments. But having someone such as you visit could be suitable for her."

Lena heard her name mentioned and turned to Paul. "He's a chess player," Jarvis said.

Lena's green eyes appraised Paul. She was about his age, he guessed. He remembered what the media had reported regarding her. She'd been in the foster-care system, had got a scholarship to college, had been majoring in theatre when the famed neuroscientist Jarvis McKnight had developed the massive staph infection that overwhelmed his bodily systems, and had volunteered for the transplant surgery that enabled the scientist to survive.



"There should be more to life than chess," Paul said. "I've been trying to become a poet."

"Tinkle tinkle little star," Lena said.

He wondered if she'd deliberately got the words wrong. "But I'll probably need a reliable career," he continued.

A complex expression flitted across her face. Was it a mixture of embarrassment and defiance? "We do what we must," she said. She turned back to the man she'd been chatting with.

Paul and Jarvis looked at each other.

"Yes," Jarvis said. "We do indeed do what we must."

Paul wanted to say to him, *I've been on the fringes too. We have that bond between us.* But he was silent.

"My research goals were too important," Jarvis said. "I couldn't let them die with me. And now I have a new goal. Can the brain be preserved against ageing if the host body that supports it is young and healthy? Perhaps cannulas bored through the skull into the brain can deliver a nourishing mixture of nutrients and antioxidants."

"And in fact why stop there?" Jarvis continued. "Those same tubes could deliver stimulants, which could improve the mind's talents. Perhaps with such injections

you could become one of the top chess players. Or would you rather become obsessed with writing poetry?"

The choice hadn't occurred to Paul as one he might have to make.

"A few volunteers such as yourself could change the world," Jarvis said. "Maybe."

Before Paul got a chance to reply, a woman rushed up. "I've so wanted to meet you," she said in a gushing voice as she held out her hand to Jarvis. Then, realizing she'd committed a social blunder, she put her hand down and blushed.

"I'd shake your hand if I could," Jarvis said in a kindly tone.

"I've admired your courage, your commitment, your dedication to your ideals," she said.

"I think of it as desperation, instead," Jarvis said. "But thank you."

The woman retreated.

"We humans have our backs against the wall," Jarvis said quietly. "Perhaps some people don't realize that yet."

Paul didn't know what to say. Then he decided to make a jesting comment. "I'll cry if you say things like that."

Jarvis said simply, "I've cried, too." Then he suddenly affixed his eyes on Paul's. "Would you like to have tubes

drilled into your brain? Would you like new biochemical substrates that could give you a greatly expanded repertoire of abilities?"

At that point Lena turned her head toward Jarvis. Paul saw her companion had vanished. "I think you might be seeing more of Roberto," she said to Jarvis. Then she spoke to Paul. "Jarvis and I have an understanding. He doesn't expect me to be celibate. We discussed this before the operation."

"I'll bite it off if I get the chance," Jarvis said.

She laughed nervously. "Jarvis has an exquisite sense of humour. The two of us might go far together. I've been wondering if, in time, we might run for president. He with his intellect and I with my theatrical skills ... What great leadership we could provide."

All in all, Paul decided later, it had been an interesting evening.

Then he realized he hadn't answered Jarvis's question. ■

Dan Gollub is a psychologist. He had a story published in *Writers Of The Future, Volume X*. He has an original approach to dream interpretation which can be found at thenakedscientists.com.

JACEY